IMBEDDING THEOREMS IN DIFFERENTIAL GEOMETRY*

T. Y. THOMAS

The matters which I shall discuss today will be largely concerned with the general problem of imbedding a coordinate manifold of class C^r in a euclidean space of sufficiently many dimensions. I imagine you all have a fair idea as to what is meant by an n-dimensional manifold of class C^r. Briefly this may be described as a Hausdorff space each point of which admits a neighborhood homeomorphic to the interior of a sphere in an n-dimensional euclidean space. We then suppose that the coordinate systems which may be introduced into these neighborhoods by these homeomorphisms are such that the coordinate relations which exist in the intersection of two such coordinate neighborhoods are of class C^r, that is, possess continuous partial derivatives to the order r inclusive. Of course when we consider imbedding theorems in differential geometry, which is the title of this address, one usually thinks, possibly from historical reasons incidental to the development of the subject, that the given space is endowed with a Riemann metric, and is then concerned with the problem of isometric imbedding. We shall also have something to say about the problem of isometric imbedding, although in doing so we shall limit ourselves to results of a general character. Of necessity most of the mathematical details must be omitted from our discussion however interesting these may be, but here and there certain detailed considerations will be introduced when it appears that these are directly understandable and may be treated with dispatch.

Before proceeding to the discussion of our particular subject I should like to say a few words about the analogous purely topological imbedding problem. I have in mind principally the classical result of Menger [1] and Nöbeling [2] to the effect that every n-dimensional compact metric space is homeomorphic to a subset of the euclidean space of $2n+1$ dimensions. Let A and B be two compact metric spaces and let M denote the set of all continuous maps of A into subsets of B. If f and f' are two elements or points of M, we define the distance between these points to be the maximum value of the distance of the points $f(x)$ and $f'(x)$ as x runs over the points of the space A. With this definition of distance M becomes a metric space and may readily be shown to be complete, that is, every Cauchy sequence in M con-

* An address delivered before the Stanford meeting of the Society on April 15, 1939, by invitation of the Program Committee.

verges. Now Hurewicz [3] has given a very simple and elegant proof of the Menger-Nöbeling imbedding theorem based on his idea of the ϵ-map. An element f of the above metric space M is called an ϵ-map of A into B if every two points of A whose map points coincide in B have a distance less than ϵ from one another. Evidently if f is an ϵ-map for every $\epsilon>0$, it is a homeomorphic map of A into B and the intersection of any sequence of ϵ-maps with ϵ approaching zero is identical with the set of all homeomorphic maps of A into B. Now it may be shown that for any $\epsilon>0$ the ϵ-maps form an open set in M. Hurewicz then proceeds to show that if B is taken to be a complete sphere in the euclidean space of $2n+1$ dimensions, the set of all ϵ-maps for any value of $\epsilon>0$ is dense in the map space M. It follows by the Baire theorem that the compact metric space A can be mapped topologically into a subset of the $2n+1$ dimensional euclidean space and in fact that any continuous map of A into this euclidean space can be changed into a topological map by an arbitrarily small alteration.

I should now like to turn to the consideration of the work of Hassler Whitney [4] on the imbedding of coordinate manifolds in euclidean space. Owing to the extreme elegance of the treatment of the purely topological imbedding problem the methods there employed may very well serve as a model or perhaps as a goal toward which we may strive in our discussion of the imbedding of coordinate manifolds. I propose therefore to present some of the results of Whitney from this standpoint. First let me say that if the question is raised as to why the topological imbedding theorem does not suffice in the present case, the answer is the following: In the imbedding problem for coordinate manifolds of class C^r we demand more. We require in fact that the functions defining the imbedding shall be of class C^r and shall thus leave unaffected the underlying coordinate relationships which we recognize as a component part of the structure of the manifold.

In accordance with the above proposal let us impose the condition that our coordinate manifold is compact and metrizable. Now we know from point set theory that a compact Hausdorff space can be given a metric in the topological sense if, and only if, it is separable. We shall now show that a positive definite quadratic differential form can be defined over a compact and separable manifold of class C^r, the coefficients of this form being of class C^{r-1} as functions of the allowable coordinates of the manifold. Thus the coordinate manifold becomes a Riemann space of class C^{r-1} in the usual terminology.

It is easy to show the existence of a function $h(x)$ of class C^∞ de-

fined for $-\infty < x < \infty$ such that

(a) $h(x) = 0, \qquad x \leqq -1, x \geqq 1,$

(b) $h(x) = 1, \qquad -1/2 \leqq x \leqq 1/2,$

(c) $0 < h(x) < 1, \qquad -1 < x < -1/2, 1/2 < x < 1.$

Now let P be any point of our coordinate manifold and denote by $N(P)$ a coordinate neighborhood of P. We suppose the coordinates x^α so chosen in $N(P)$ that $x^\alpha = 0$ at P. By a coordinate transformation in $N(P)$ of the form $\bar{x}^\alpha = ax^\alpha$ (a = const.) we can so enlarge the coordinate representation of this neighborhood that it will contain a cube or box $-1 \leqq x^\alpha \leqq 1$. Denote the interior $(-1 < x^\alpha < 1)$ of this box by U and the box itself or the enclosure of U by $\bar{U}$. We shall also consider the smaller box $\bar{U}'$ defined by $-1/2 \leqq x^\alpha \leqq 1/2$ and its interior U'. In consequence of the above assumptions (compactness and separability) the coordinate manifold M is bicompact, that is, every covering of the manifold by open sets contains a finite covering. Hence a covering of the manifold M by the above open sets U' will contain a finite covering which we shall denote by $U_1', \cdots, U_t'$. Let us now put

$$H = h(x^1) \cdots h(x^n) \text{ in } \bar{U}, \qquad H = 0 \text{ in } M - \bar{U},$$

$$H_\alpha = x^\alpha H \text{ in } \bar{U}, \qquad H_\alpha = 0 \text{ in } M - \bar{U},$$

where U refers to any one of the coverings of the above finite set. Then $H = 1$ and $H_\alpha = x^\alpha$ in $\bar{U}'$. It is evident that the functions H and H_α are scalars of class C^r over M. Let us denote the functions H_α when determined in connection with the neighborhood U_1' by $\phi^1, \cdots, \phi^n$ and when determined in connection with the neighborhood U_2' by $\phi^{n+1}, \cdots, \phi^{2n}$, and so on. Then the equations

$$g_{\alpha\beta} = \sum_\nu \frac{\partial \phi^\nu}{\partial x^\alpha} \frac{\partial \phi^\nu}{\partial x^\beta}$$

define the components of a covariant tensor g of class C^{r-1} over M. Now the matrix $\|\partial\phi^\nu/\partial x^\alpha\|$ has rank n at every point P of M. In fact any point P is interior to one of the neighborhoods U_i' and in this neighborhood $H_\alpha = x^\alpha$. Hence the matrix $\|\partial\phi^\nu/\partial x^\alpha\|$ must contain the diagonal determinant $|\delta_\beta^\alpha|$ at P. Thus at any point P of M the above functions $g_{\alpha\beta}$ appear as the coefficients of a positive definite quadratic differential form by which a Riemann metric is defined in M.

Conversely if a (connected) compact manifold M admits as above a Riemann metric g, we may metrize M in the topological sense by defining the distance $\rho(P, Q)$ of two points P and Q as the greatest lower bound of the lengths of all curves of class C^r (broken or con-

tinuously differentiable) which join P to Q. It follows that M is separable. The result obtained may perhaps be stated as the following theorem.

RIEMANN METRIZATION THEOREM. *Any compact manifold M of class C^r with $r \geqq 1$ admits a Riemann metric of class C^{r-1} if, and only if, M is separable.*

From now on we shall consider a fixed metric g of class C^{r-1} in M which exists in accordance with the above theorem. When it is desired to measure distances independently of coordinate systems this metric will be used. Within the separate coordinate systems the euclidean metric may however sometimes be employed to advantage.

Let f be a map of class C^r of M into the euclidean space E_m. The dimensionality m of the euclidean space will not be fixed for the present but will be determined later on the basis of our discussion. Let S be the set of all such maps f. We define a metric in S in the following manner: If ϕ and ψ are two elements of S, let us put

$$d_0(\phi(x), \psi(x)) = \{ \sum [\phi^i(x) - \psi^i(x)][\phi^i(x) - \psi^i(x)] \}^{1/2},$$
$$d_1(\phi(x), \psi(x)) = \{ \sum g^{\alpha\beta}[\phi^i_{;\alpha} - \psi^i_{;\alpha}][\phi^i_{;\beta} - \psi^i_{;\beta}] \}^{1/2},$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots,$$
$$d_r(\phi(x), \psi(x)) = \{ \sum g^{\alpha\mu} \cdots g^{\beta\nu}[\phi^i_{;\alpha\cdots\beta} - \psi^i_{;\alpha\cdots\beta}][\phi^i_{;\mu\cdots\nu} - \psi^i_{;\mu\cdots\nu}] \}^{1/2},$$

where it is to be understood that Latin indices have the range $1, \cdots, m$ and that Greek indices have the range $1, \cdots, n$ and both sets of indices are to be summed when repeated in accordance with the usual convention. Note also that $d_1, \cdots, d_r$ involve successive covariant derivatives of the functions ϕ and ψ and that d_r involves the rth covariant derivative which is the highest covariant derivative that can be formed under the hypothesis that the metric g is of class C^{r-1} and the maps ϕ and ψ are of class C^r. Put

$$D(\phi(x), \psi(x)) = d_0(\phi(x), \psi(x)) + \cdots + d_r(\phi(x), \psi(x)).$$

Then $D(\phi(x), \psi(x))$ is a continuous function on M. Since M is bicompact, the function $D(\phi(x), \psi(x))$ assumes its maximum value at a point of M. Denote this maximum value by $D(\phi, \psi)$ and define $D(\phi, \psi)$ as the distance between the points ϕ and ψ of S. It is easily seen that with this definition of distance S is a metric space. Moreover it can readily be shown that S is complete. The fact that the map space S whose elements are the maps of class C^r of the coordinate manifold M into the euclidean space E_m thus appears as a complete metric space is in strict analogy with the situation in the topological

imbedding problem, and imbedding theorems which are likewise analogous to those of the topological theory can be demonstrated.

We shall say that a map or point ϕ in S is *regular* if the matrix $\|\partial\phi^i/\partial x^\alpha\|$ has rank n at every point of M. Otherwise ϕ will be said to be *singular* and the points at which the above matrix has rank less than n will be called the *singular points* of the map. The set of all such singular points will be called the domain of *singularity* of the map ϕ.

As one can readily imagine, there are many details in this demonstration which have no counterpart in the purely topological theory. It seems inadvisable to enter into such details here especially in view of the fact that the main results can be stated directly in terms of the point of view which we have now established and when so stated are immediately understandable. We shall therefore content ourselves here with the following statement of what may be classified as two of the main imbedding theorems for coordinate manifolds:

I. *If $r \geqq 2$ and $m \geqq 2n$, the regular maps form an open and dense set in S.*

II. *If $r \geqq 2$ and $m \geqq 2n+1$, the regular topological maps form an open and dense set in S.*

Since $f=0$ is an element of S, it follows that the sets which enter into the above theorems are nonvacuous. Thus any compact and separable manifold of class C^r with $r \geqq 2$ can be imbedded by a regular topological map of class C^r in the euclidean space of $2n+1$ dimensions and an infinitesimal alteration in any map of class C^r will result in a regular topological map in strict analogy with the topological imbedding theorem (Whitney [4]).

There is perhaps a mild interest in considering the extreme case in which the imbedding of the coordinate manifold is in the euclidean space E_n and thus of the same dimensionality as the manifold itself. Assuming as above that the manifold is compact and separable and of $n \geqq 2$ dimensions, it suffices now to suppose that it is of class C^r with $r \geqq 1$. We now let S be the complete metric space whose elements are the maps of class C^r of the manifold M into the euclidean space E_n and denote by Σ the set of all points in S which correspond to maps whose domain of singularity is nowhere dense in M. It may then be proved that the following theorem holds.

III. *The set Σ is dense in S.*

It follows that the manifold M can be imbedded in the euclidean space of n dimensions by a map of class C^r whose domain of singularity is closed and nowhere dense in M. In particular it is thus possible

to define a Riemann metric of class C^{r-1} in M which is nonsingular and locally flat in the ordinary sense except over a closed and nowhere dense set in the space (Thomas [5]).

The methods and results which we have so far considered cannot be applied to the case of the analytic manifold. Nor shall we now have much to say about such manifolds since as far as I am aware there are no results of a general character pertaining to the imbedding of analytic manifolds in euclidean space. Recently Bochner [6] has shown that a compact and separable analytic manifold which bears an analytic Riemann metric, that is, a compact analytic Riemann space, can be imbedded analytically and topologically in a euclidean space of $2n+1$ dimensions. Bochner's demonstration consists in showing that it is possible to approximate a map of class C^r with $r \geqq 2$ of the manifold in the euclidean space as closely as desired by an analytic map. Selecting the map of class C^r to be topological, it follows that not only will the analytic map exist but it will likewise be topological. Yet it must be confessed that this result does not constitute a strict imbedding theorem for analytic manifolds owing to the assumption of the existence of the Riemann metric. In fact under this assumption one would naturally inquire into the question of the isometric imbedding (in the differential sense) of the Riemann space into euclidean space. Some time ago I considered the question of the possibility of joining two or more arbitrary points of an analytic manifold by an analytic arc and was able to accomplish this only under the assumption that the manifold possessed an analytic affine connection [7]. While the assumption of the existence of the affine connection is presumably somewhat weaker than Bochner's assumption of the Riemann metric, it is nevertheless an assumption of analogous character. Apparently one is here faced with an underlying difficulty the essential idea for the solution of which remains as yet undiscovered.

Let us now turn our attention to the problem of the isometric imbedding of a Riemann space in euclidean space to which in fact our discussion has naturally led. This problem may be identified with the problem of finding a solution $\phi_1, \cdots, \phi_m$ of the system

$$g_{\alpha\beta} = \sum_{i=1}^{m} \frac{\partial \phi_i}{\partial x^\alpha} \frac{\partial \phi_i}{\partial x^\beta}$$

over the Riemann space with fundamental metric tensor g. The result that an n dimensional Riemann space can be imbedded locally and isometrically in a euclidean space of $n(n+1)/2$ dimensions seems to have been enunciated first by Schlaefli [8]. Janet [9] is usually con-

sidered the first to have made a serious attempt to prove the local isomorphic imbedding theorem for Riemann spaces, the result at which he arrived being that above stated. This was followed by proofs of the same theorem by Cartan [10] and Burstin [11]. In all cases the theorem in question was made to depend on more or less standard results in the theory of systems of differential equations and so need not be considered further on this occasion.

With regard to the question of the isomorphic imbedding of Riemann spaces in the large in euclidean space no general result seems to be known. It would appear therefore that here one would find an interesting although difficult field for investigation.

It is of course well known that the condition for the local isomorphic imbedding of a Riemann space of n dimensions in the euclidean space of n dimensions is that the curvature tensor shall vanish over the Riemann space. If furthermore the given Riemann space is simply connected, this condition suffices for its complete imbedding in the n dimensional euclidean space.

The problem of determining the conditions under which an n dimensional Riemann space is of class one, that is, can be imbedded isometrically in a euclidean space of $n+1$ dimensions but not in an n dimensional euclidean space, admits in general a solution of some degree of refinement. As I have occasion to believe that this imbedding theorem is not generally well known I should like to indicate in slight detail at least some of the essential features on which the solution of this problem depends.

Let R denote a Riemann space of class C^2. It is then well known that the following system of equations,

$$\frac{\partial y^i}{\partial x^\alpha} = y^i_{,\alpha}, \quad \frac{\partial y^i_{,\alpha}}{\partial x^\beta} = \sum_{\nu=1}^{n} \Gamma^\nu_{\alpha\beta} y^i_{,\nu} + b_{\alpha\beta}\sigma^i, \quad \frac{\partial \sigma^i}{\partial x^\alpha} = \sum_{\mu,\nu=1}^{n} b_{\alpha\mu} g^{\mu\nu} y^i_{,\nu},$$

constitutes necessary conditions for the isomorphic imbedding of R in $n+1$ dimensional euclidean space, that is, for R to appear as a hypersurface in the euclidean space. In these equations the y^i denote the coordinates of the euclidean space, the Γ's are the components of the Christoffel symbols of the Riemann space R, the b's are the coefficients of the second fundamental form of the hypersurface and the σ's are the components of the vectors in the euclidean space normal to the hypersurface. As integrability conditions of the above system we have

$$b_{\alpha\beta,\gamma} = b_{\alpha\gamma,\beta} \qquad \text{(Codazzi equations)},$$

$$B_{\alpha\beta\gamma\delta} = b_{\alpha\delta} b_{\beta\gamma} - b_{\alpha\gamma} b_{\beta\delta} \qquad \text{(Gauss equations)},$$

in which the first set of these equations involves the components of the covariant derivative of the above tensor b and the second set of equations contains in its left member the components of the completely covariant form of the curvature tensor of the Riemann space R. In the above general form these equations were first obtained by Voss. They are usually however referred to as the Gauss and Codazzi equations since they are equivalent to conditions originally found by Gauss and Codazzi for the special case of two dimensional surfaces. The usual isomorphic imbedding theorem for Riemann spaces of class one can now be stated as follows: *An open and simply connected coordinate neighborhood U of the above Riemann space R can be imbedded isomorphically in the $n+1$ dimensional euclidean space if, and only if, the Gauss and Codazzi equations are satisfied in U.* It is well known too that the set of quantities $b_{\alpha\beta}$ satisfying the Gauss and Codazzi equations in the neighborhood U appear as the coefficients of the second fundamental form of the hypersurface which exists by the above theorem and that when these quantities $b_{\alpha\beta}$ are fixed the hypersurface is determined to within a motion in the euclidean space.

The above theorem cannot be considered to give a solution of the local isometric imbedding problem in any fundamental sense. For the conditions in question are of differential character and in this respect are analogous to the conditions by which the imbedding itself is defined. On the other hand the condition for a Riemann space to be locally flat or to be capable of being imbedded locally in a euclidean space of the same dimensionality is expressible by the vanishing of a pure invariant, namely the curvature tensor of the given Riemann space. It is likewise possible to express conditions for a Riemann space to be of class one in terms of the behavior of its intrinsic invariants and I should now like to indicate some of the steps by which these conditions can be established.

Let us say that a hypersurface of the euclidean space of $n+1$ dimensions is of type one if the rank of the matrix of the coefficients of the second fundamental form is zero or one and that the hypersurface is of type τ where τ is an integer of the set $2, \cdots, n$, if the rank of this matrix is τ over the hypersurface. As so defined the type number is not an intrinsic invariant but depends upon the relation of the space to the euclidean space in which it is imbedded. It can be shown however that the type number of a hypersurface is determined by its intrinsic properties, that is, by the first fundamental form. As a consequence we can speak of the type number of a Riemann space R of class C^2 regardless of whether or not this space can be considered as a hypersurface of the euclidean space. Let us also say that a hypersur-

face is *intrinsically rigid* if the second fundamental form is uniquely determined (to within algebraic sign) by the first fundamental form and the equations of Gauss and Codazzi. It follows that an intrinsically rigid hypersurface cannot be subjected to a continuous deformation in the euclidean space without altering its internal metric properties. The following theorem can also be proved: *A hypersurface of type* $\tau \geqq 3$ *is intrinsically rigid.*

The solution of the problem before us depends essentially on the happy circumstance that under certain rather general conditions the equations of Codazzi are consequences of the equations of Gauss. In fact it can be shown that if R is any Riemann space of class C^2 and type not less than 4 and if there exists a set of symmetric quantities $b_{\alpha\beta}$ of class C^1 in a coordinate neighborhood U which satisfy the equations of Gauss, then the equations of Codazzi will automatically be satisfied in U. With this result the determination of the conditions for a Riemann space to be of class one is reduced essentially to an algebraic problem. It can be shown that the Gauss equations considered as equations for the determination of the symmetric quantities $b_{\alpha\beta}$ admit a resultant system and to this system further algebraic conditions can be added in the form of inequalities which are both necessary and sufficient for the reality of the solutions. Assuming that the Riemann space under consideration is of class C^2 and that the components of its curvature tensor are continuous functions, which is somewhat weaker than the requirement that the space be of class C^3, the general result at which one arrives is the following: *There exist sets of polynomials* F_1, F_2 *and* F_3 *in the components of the curvature tensor such that a simply connected* n*-dimensional Riemann space of type not less than* 4 *can be imbedded isometrically in the euclidean space of* $n+1$ *dimensions if, and only if,* $F_1 > 0$, $F_2 \geqq 0$ *and* $F_3 = 0$ *over the space* (Thomas [12]). For want of a better name I have said that such conditions constitute an *algebraic characterization.* With some modifications in procedure it is possible to extend the above result to Riemann spaces of type 3 and in fact to spaces of variable type. Moreover the type of a Riemann space is itself capable of being characterized algebraically and so may be included if desired in the algebraic characterization. It must be emphasized however that without the above restriction on the type of the Riemann space these algebraic procedures are not applicable and indeed it is questionable if the class of Riemann spaces which can be imbedded isometrically in the euclidean space of one higher dimension admits an algebraic characterization.

These results have recently been extended by C. B. Allendoerfer [13] to Riemann spaces of class greater than one although the meth-

ods exhibit considerable formal complication owing without doubt to the nature of the problem.

References

1. Menger, *Dimensionstheorie*, Leipzig, 1928, p. 296.
2. Nöbeling, Mathematische Annalen, vol. 104, p. 71.
3. Hurewicz, Sitzungsberichte der preussischen Akademie der Wissenschaften, Berlin, Physikalisch-mathematische Klasse, 1933.
4. Whitney, Annals of Mathematics, (2), vol. 37 (1936), p. 645.
5. Thomas, American Journal of Mathematics, vol. 61 (1939), p. 403.
6. Bochner, Duke Mathematical Journal, vol. 3 (1937), p. 339.
7. Thomas, Annals of Mathematics, (2), vol. 38 (1937), p. 120.
8. Schlaefli, Annali di Matematica Pura ed Applicata, (2), vol. 5 (1871–1873), p. 170.
9. Janet, Annales de la Société Polonaise de Mathématique, vol. 5 (1926), p. 38.
10. Cartan, Annales de la Société Polonaise de Mathématique, vol. 6 (1927), p. 1.
11. Burstin, Receuil Mathématique, Moscow, vol. 38, p. 74.
12. Thomas, Acta Mathematica, vol. 67 (1936), p. 169.
13. Allendoerfer, Duke Mathematical Journal, vol. 3 (1937), p. 317.

University of California at Los Angeles

THE DIRICHLET PROBLEM FOR THE VIBRATING STRING EQUATION

D. G. BOURGIN AND R. DUFFIN

This note considers the Dirichlet and Neumann type boundary value problem for the simple vibrating string equation. The detailed study for a special boundary is timely in view of certain categorical statements in the recent literature.* The results obtained below indicate how such statements are to be modified.† Of independent interest is the novel procedure, stemming from Lemma 1, for proving uniqueness in Theorems 1 and 2. The method is of wide utility and leads to interesting generalizations.

For convenience we use t for $v\tau$, where v and τ refer to the velocity of wave propagation and the time, respectively. The string equation is then

$$L[y] \equiv y_{xx} - y_{tt} = 0, \tag{1}$$

and the data are given on the boundaries of the finite rectangle

$$0 \leqq x \leqq S, \qquad 0 \leqq t \leqq T. \tag{1.1}$$

We denote the ratio T/S by α. The term "rectangle," used in the sequel, unless otherwise qualified, refers to the *closed* rectangle defined in (1.1).

We shall need the following lemma.

LEMMA 1. *If $y(x, t)$ is continuous in both real arguments in the rectangle, and if $F(p, u)$ is defined as*

$$F(p, u) = \int_0^T \int_0^S e^{i(xp+tu)} y(x, t)\, dx dt, \tag{2}$$

then $F(p, u)$ is an entire function in each of the complex variables p and u.

* For example: "Dagegen würde ein Randwertproblem im Falle unserer hyperbolischen Differentialgleichung (our (1)) sinnlos sein," Courant-Hilbert, *Methoden der Mathematischen Physik*, vol. 2, p. 178. ". . . le problème de Dirichlet ne peut se poser pour le cas hyperbolique," J. Hadamard, L'Enseignement Mathématique, vol. 35 (1936), p. 26. Some of Hadamard's surmises are not borne out for the special situation we treat. Cf. pp. 26, 29 and notes on p. 29, loc. cit.

† In "principle," a physical realization of the Dirichlet problem is afforded by taking photographs of a vibrating string at two different times. However, "practically," our analysis is entirely ineffective, not for the reason of overdetermination (as in the illuminating instance on which Courant founds his mathematical conclusion) but because of the physical unpreciseness and inconstancy of the all important α.

The lemma is easily established even for less restricted $y(x, t)$.*

THEOREM 1. *Let $y(x, t)$ be a solution of* (1) *in the interior of the rectangle and let $y(x, t)$ vanish on the boundary. Furthermore, $y(x, t)$ is to be of class C'† in the closed rectangle and y_{xx} and y_{tt} are to be Lebesgue summable over the rectangle. Then, in order that the solution $y(x, t) = 0$ be unique, it is necessary and sufficient that α be an irrational number.*

We take up the "sufficiency" aspect first. Our hypotheses justify the application of Green's theorem‡ to (1) in the form

$$\int_0^T \int_0^S \left\{e^{i(px+ut)}L[y] - yL[e^{i(px+ut)}]\right\} dxdt$$

$$(3) \qquad = (p^2 - u^2)F(p, u) = \int_0^S (e^{i(px+ut)}(y_t(x, t) - iuy(x, t)))\Big|_{t=0}^{t=T} dx$$

$$- \int_0^T (e^{i(px+ut)}(y_x(x, t) - ipy(x, t)))\Big|_{x=0}^{x=S} dt.$$

On setting $y = 0$ on the boundary, we obtain in the notation of (2)

$$(p^2 - u^2)F(p, u)$$

$$(3.1) \qquad = \int_0^S e^{i(px+ut)}y_t(x, t)\Big|_{t=0}^{t=T} dx - \int_0^T e^{i(px+ut)}y_x(x, t)\Big|_{x=0}^{x=S} dx.$$

Since $F(p, u)$ is entire, according to Lemma (1), the left side of (3.1) vanishes for $p = \pm u$. Therefore

$$(3.2) \qquad \int_0^S (e^{iu(t+x)}y_t(x, t))\Big|_{t=0}^{t=T} dx - \int_0^T (e^{iu(t+x)}y_x(x, t))\Big|_{x=0}^{x=S} dt = 0,$$

$$(3.21) \qquad \int_0^S (e^{iu(t-x)}y_t(x, t))\Big|_{t=0}^{t=T} dx - \int_0^T (e^{iu(t-x)}y_x(x, t))\Big|_{x=0}^{x=S} dt = 0.$$

On subtracting these equations and taking the real part, there results

$$(3.3) \qquad \sin uT \int_0^S y_t(x, T) \sin ux dx - \sin uS \int_0^T y_x(S, t) \sin ut dt = 0.$$

In view of the incommensurability of T and S, the special choices $u = n\pi/T$ and $u = n\pi/S$ imply

---

* Theorem 5.31, Whittaker and Watson, *Modern Analysis*, 3d edition, is the main tool.

† Class C^n implies continuity in all arguments through derivatives of order n.

‡ The essentials of the argument are given in Horn, *Partielle Differentialgleichungen*, 2d edition, p. 107.

$$(3.4) \qquad \int_0^T y_x(S, t) \sin n\pi t/T \, dt = \int_0^S y_t(x, T) \sin n\pi x/S \, dx = 0.$$

We may extend $y_t(x, T)$ to all values of x by requiring it to be an odd periodic function of period $2S$. This extended function is at worst piecewise continuous and summable and accordingly admits of a Fourier sine expansion, summable C^1.* According to (3.4) the Fourier coefficients vanish. The supposed continuity in the range $0 \leqq x \leqq S$ allows the deduction $y_t(x, T) = 0$.† Similarly $y_x(S, t) = 0$. The symmetry of the problem then guarantees the vanishing of $y_t(x, 0)$ and $y_x(0, t)$.‡ It follows from (3.1) that

$$(3.5) \qquad F(p, u) = 0$$

if $p \neq \pm u$. This latter restriction may be waived, however. In fact, consider an arbitrary value of u and denote it by u_0. $F(p, u_0)$ vanishes for all points in any open neighborhood of $p = u_0$ in the complex p plane, except, possibly, for $p = u_0$. However, since $F(p, u_0)$ is analytic in p, it follows that $F(u_0, u_0)$ is also 0. Similarly $F(-u_0, u_0) = 0$. Hence the deduction (3.5) is valid for all finite p and u. For our present purpose we may consider $y(x, t)$ to be 0 for all points outside the fundamental rectangle. Then $F(p, u)$ is an identically vanishing Fourier transform of $y(x, t)$ whence $y(x, t) = 0$.§

The necessity is easily verified. Indeed if α is rational we have the nonidentically vanishing solution

$$y(x, t) = \sin n\pi x/S \sin n\pi t/S,$$

where n and $n\alpha$ are integers.

For the case envisaged in Theorem 1 it is worthwhile to remark an alternative proof which, though not directly capable of extension to a more general equation, may clarify the physical situation. Our starting point is the classic solution of (1), namely

$$(4) \qquad 2y(x, t) = y(x + t, 0) + y(x - t, 0) + g(x + t) - g(x - t),$$

* E. C. Titchmarsh, *Theory of Functions*, pp. 414 and 426.

† Since we now have established vanishing Cauchy data on the upper side of the rectangle we can appeal to the theory of the so-called "mixed" problem to gain the conclusion in our theorem. However, it is of interest to carry through the method predicated on Lemma 1 to show its power and completeness.

‡ Alternatively we may take the real part of the difference of equations (3.2) and (3.21), and also the imaginary part of their sum. With the justifiable definition of $y_t(x, 0)$ and $y_x(0, t)$ as odd periodic functions of x and t, with periods $2S$ and $2T$ respectively, one then easily derives $y_t(x, 0) = y_x(0, t) = 0$ as a consequence of $y_t(x, T) = y_x(S, t) = 0$.

§ S. Bochner, *Vorlesungen über Fourier Integralen*, p. 192.

where

$$g(z) = \int_0^z y_t(x, 0)dx.$$

Now $y_t(x, 0)$ is of class C^0 and may be considered odd and periodic on the x axis with period $2S$. Since $g(z)$ is the integral of an odd periodic function it is itself periodic with the same period $2S$.

The identical vanishing of y on the boundary implies

$$0 = g(x + T) - g(x - T). \tag{4.1}$$

Thus $g(z)$ has a second period, $2T$. Accordingly

$$g(2nT + 2mS) = g(0) \tag{4.2}$$

for all integer values of $|n|$ and $|m|$. Since T and S are incommensurable it follows that, for arbitrary x and ϵ, positive or negative integers $\bar{n}$ and $\bar{m}$ may be found such that

$$|x + 2\bar{n}T + 2\bar{m}S| < \epsilon. \tag{4.3}$$

Relation (4.3) is, in fact, an immediate consequence of the well known Kronecker inequality. Thus the points $\{2nT+2Sm\}$ are everywhere dense on the interval

$$-S - T \leqq z \leqq S + T. \tag{4.4}$$

Manifestly $g(z)$ is continuous on the interval defined in (4.4). Since a continuous function on an interval is determined uniquely by its values on a dense set of points, it follows from (4.2) that

$$g(z) = g(0) \tag{4.5}$$

for all z in the interval in question. It is now seen that (4) guarantees $y(x, t) = 0$.

We may also establish uniqueness in the case of the Neumann problem (or more generally, for the case that Dirichlet data are given on some sides and Neumann data on the others). For instance we have the following result.

THEOREM 2. *Let $y(x, t)$ be a solution of* (1) *for interior points and be of class C^1 in the rectangle, with y_{xx} and y_{tt} summable over the rectangle. Furthermore, let y_t and y_x vanish on the horizontal and vertical sides of the boundary, respectively. Then irrationality of α is the necessary and sufficient condition that the only solution satisfying our hypotheses be a constant.*

The necessity follows from the observation that, for rational α,

$$y(x, t) = \cos n\pi x/S \cos n\pi t/S$$

is a solution for values of n such that $n\alpha$ is also an integer.

The demonstration of sufficiency requires, merely, obvious modification of the analysis in our first method of proof of Theorem 1. Thus, on setting $p = \pm u$ in (3) we obtain as the analogues of relations (3.2) and (3.21),

$$(5.1) \qquad \int_0^S e^{iu(x+t)} y(x, t) \Big|_{t=0}^{t=T} dx = \int_0^T e^{iu(x+t)} y(x, t) \Big|_{x=0}^{x=S} dt,$$

$$(5.11) \qquad \int_0^S e^{iu(x-t)} y(x, t) \Big|_{t=0}^{t=T} dx = \int_0^T e^{iu(x-t)} y(x, t) \Big|_{x=0}^{x=S} dt.$$

Manifestly we may take $y(x, T), y(x, 0), y(S, t), y(0, t)$ as even periodic functions of periods $2S$ and $2T$ according as the variable argument is x or t. Formulas (5.1) and (5.11) imply

$$(5.2) \qquad \sin uT \int_0^S \cos ux\, y(x, T)dx = \sin uS \int_0^T \cos ut\, y(S, t)dt.$$

The choices $u = n\pi/S$ and $u = n\pi/T$ show that the integrals on either side vanish for all *positive* integers n and are unrestricted for $n = 0$. We may paraphrase an earlier argument to deduce that $y(x, t)$ is a constant K on the boundary. On evaluation of the boundary integrals, (3) now yields

$$F(p, u) = K(e^{iuT} - 1)(e^{ipS} - 1)/up.$$

This is the transform of a function vanishing outside the rectangle. Inside and on the boundary

$$(5.3) \qquad y(x, t) = K.$$

The restriction that y be of class C^1, coupled with the almost everywhere equivalence of two Lebesgue summable functions having the same transform, implies that (5.4) is unique.

THEOREM 3. *Let $y(x, t)$ satisfy all the conditions imposed in Theorem 1 save that the boundary values are now arbitrary functions of class C^1 which vanish at the four corners. Then if α is irrational, a solution $y(x, t)$, if it exists, is unique.*

The demonstration is along conventional lines. Suppose y_1 and y_2 are distinct solutions. Define y^* by $y^* = y_1 - y_2$. Clearly y^* satisfies the hypotheses of Theorem 1 whence it follows that y_1 and y_2 are identical.

The ratio α plays a determining role in the demonstration of actual existence of solutions. Our attention will be focussed on α values satisfying the inequality (E)

$$|\alpha - m/n| > A/n^{K+1}, \tag{E}$$

where A and K are fixed, m, n, K are positive integers and $m \leqq 2\alpha n$. The algebraic numbers of degree $K+1$ satisfy (E). Moreover the set of α values which do not satisfy (E) for any choice of K is known to be of linear Lebesgue measure zero.*

LEMMA 2. *If α satisfies* (E), *then for each integer N, sufficiently large, there is associated an integer M', depending on N, such that* $1/2N > |1/\alpha - M'/N| > B/N^{K+1}$.

We replace n by M and m by N in (E). Clearly we may select a "best" M, denoted by M', for each N in the sense

$$|\alpha - N/M| \geqq |\alpha - N/M'|, \qquad M \neq M'. \tag{6}$$

For N_0 sufficiently large and fixed

$$2\alpha > N/M' > \alpha/2, \qquad N > N_0. \tag{6.1}$$

Hence by (E) and the defining relation (6) for M', we plainly have

$$\begin{aligned} 1/2N \geqq |1/\alpha - M'/N| &= |\alpha - N/M'|\, M'/\alpha N \\ &\geqq \frac{A}{\alpha}\left(\frac{N}{M'}\right)^K \Big/ N^{K+1} = B/N^{K+1}, \qquad N > N_0, \end{aligned} \tag{6.2}$$

where $B = (A/\alpha)(\alpha/2)^K$.

The assigned boundary functions $y(x, 0)$, $y(x, T)$ on the two horizontal sides of the rectangle of (1) may be extended to all x values by requiring the extended functions to be odd and periodic of period $2S$. Similarly the extended functions $y(0, t)$, $y(S, t)$ are odd and periodic of period $2T$. It is in this sense that subsequent references to "extended boundary functions" are to be understood.

THEOREM 4. *If* (a) *α satisfies* (E) *for a fixed K and if* (b) *the four extended boundary functions are of class*† *C^{4+K} for $-S \leqq x \leqq S$ or $-T \leqq t \leqq T$, then there exists a unique solution $y(x, t)$ of* (1), *satisfying the boundary conditions, such that* (c) *y_{xx} and y_{tt} are summable over the rectangle and* (d) *y is of class C^1 in the rectangle. This solution is actually of class C^2 in the rectangle.*

* E. Borel, *Théorie des Fonctions*, 3d edition, p. 27.

† The condition on the $(K+4)$th derivatives may evidently be weakened.

Condition (b) implies that the boundary functions may be expanded in Fourier sine series. Furthermore

$$a_n, A_n, b_n, B_n = o(n^{-4-K}),^* \tag{7}$$

where the constants on the left side of (7) are the nth coefficients in the sine series expansions of $y(x, 0)$, $y(x, T)$, $y(0, t)$ and $y(S, t)$, respectively. For any n we may find m' depending on n, such that

$$|\alpha n - m'| < 1/2. \tag{7.1}$$

Then for $n \geqq n_0 \geqq \max(3, A^{1/K})$,

$$|\sin n\pi\alpha| = |\sin \pi(n\alpha - m')| \geqq |\sin \pi A/n^K| \geqq A/4n^K. \tag{7.2}$$

In view of Lemma 2 we may establish in a similar way

$$|\sin n\pi/\alpha| \geqq B/4n^K, \qquad n \geqq n_1. \tag{7.3}$$

We assert that a solution satisfying the conditions of our theorem is given by

$$y(x, t) = \sum \left\{ \frac{a_n \sin (n\pi(T - t)/S) + A_n \sin n\pi t/S}{\sin n\pi\alpha} \sin n\pi x/S \right. \\ \left. + \frac{b_n \sin (n\pi(S - x)/T) + B_n \sin (n\pi x/T)}{\sin n\pi/\alpha} \sin n\pi t/T \right\}. \tag{7.4}$$

The right side of (7.4) formally satisfies (1). In view of (7), (7.2), and (7.3) it is clear that

$$\left| \frac{a_n}{\sin n\pi\alpha} \right|, \cdots, \left| \frac{B_n}{\sin n\pi/\alpha} \right| = o(n^{-4}). \tag{7.5}$$

This implies that the right side of (7.4) is the sum of four uniformly convergent Fourier series and these series remain uniformly convergent on differentiation twice by either x or t. Accordingly, the four formal series on the right side of (7.4) are Fourier series defining functions certainly of class C^2 in the rectangle. Thus y_{xx}, y_{tt} are summable. Finally, no other solution of our problem of class C^1 with summable y_{xx} and y_{tt} in the rectangle can exist, according to Theorem 2.

Incidentally, the form of (7.4) indicates that if α is rational, then a necessary condition for the existence of a (non-unique) solution is that $A_\nu = a_\nu \cos \nu\pi\alpha$, $B_\mu = b_\mu \cos \mu\pi/\alpha$, where ν and μ are integers such that $\nu\alpha$ and μ/α are also integers.

Hypothesis (b) of Theorem 4 is a *sufficient* but by no means *necessary* condition. Consider for instance, for $\alpha < 1$,

* Titchmarsh, loc. cit.

$$(8)\quad \begin{aligned} y(x,t) &= x(T-S-t)(x^2+(T-t)^2-2S(T-t));\ 0 \leqq x \leqq T-t, \\ &= (x-S)(T-t)(x^2+(T-t)^2-2Sx);\ T-t \leqq x \leqq S. \end{aligned}$$

Independently of the value of α, $y(x, 0)$ is merely of class C^2, yet $y(x, t)$ is a solution of class C^2.

Nevertheless some such hypothesis as (b) must be admitted to guarantee existence of the required solution in all cases. We take $A=1$ and denote by $K'+1$ the smallest integer K for which α satisfies (E) except, possibly, for a finite number of values of n. We shall suppose $K' \geqq 3$ in the sequel. Accordingly there exists an infinite number of pairs of values of n and m, denoted by n_i and m_i, such that

$$(8.1)\qquad |\alpha - m_i/n_i| < 1/n_i^{K'+1}.$$

Hence for $n_i > 3$

$$(8.2)\qquad |\sin n_i\pi\alpha| \leqq \pi/n_i^{K'},\ |\cot n_i\pi\alpha| \geqq n_i^{K'}/2\pi.$$

We take Dirichlet data vanishing except on the base and the following coefficients, a_n, in the Fourier sine expansion of $y(x, 0)$:

$$(8.3)\qquad a_{n_i} = 1/n_i^{K'+1},\ a_n = 0, \qquad n \neq n_i.$$

Thus the extended $y(x, 0)$ is certainly of class $C^{K'-1}$. Suppose a solution $y(x, t)$ of class C^2 exists. It is easily shown then by an argument based on (4) for instance, if we bear in mind the prescribed vanishing of $y(x, t)$ on the vertical sides, that $y_t(x, 0)$ may be extended to an odd periodic function of class C^0. Moreover $y_t(x, 0)$ is of class C^1 on the base of the rectangle. Accordingly $y(x, t)$ and $y_t(x, 0)$ are expansible in absolutely and uniformly convergent Fourier series lacking the terms $\cos \pi nx/S$. The nonzero Fourier coefficients, a_n' for $y_t(x, 0)$ must be

$$(8.4)\qquad a_{n_i}' = -\pi(\cot n_i\pi\alpha)/Sn_i^{K'}.$$

In view of (8.2)

$$(8.5)\qquad |a_{n_i}'| > 1/2S.$$

This contradicts the deduction that the $\{a_n'\}$ are Fourier coefficients. Hence $y(x, t)$ cannot be of class C^2 in the rectangle. Therefore hypothesis (b) cannot be weakened to admit data of continuity class inferior to $C^{K'}$ and, indeed, if $K' \geqq 4$ an obvious extension shows that the lowest class is $C^{K'+1}$ at least.

UNIVERSITY OF ILLINOIS AND
PURDUE UNIVERSITY

THE LAPLACE HEAVISIDE METHOD FOR BOUNDARY VALUE PROBLEMS*

D. G. BOURGIN AND R. DUFFIN

The usual methods for solution of boundary value problems for linear differential equations are based on taking linear combinations with "undetermined" coefficients of particular solutions of the homogeneous or nonhomogeneous equation. Subsidiary calculations are then required to deduce coefficient values such that the boundary conditions are satisfied. This note is concerned primarily with the formulation of a novel viewpoint in the Laplace integral treatment of differential equations and presents two new methods, neither involving undetermined coefficients, which are alternatives to the standard procedures.

Our developments stem from relations (α). The Laplace transform $F(p)$ plays the determining role. This function does not appear in the current developments of the definite integral solutions for variable coefficient equations. Since our methods yield solutions automatically satisfying prescribed boundary relations, they bear the same relationship to other procedures as does the Heaviside method, heretofore available for the one point problem for constant coefficient equations. Indeed our Method 1 is shown to be the natural generalization of the Heaviside procedure.

Consider, for real a, b,

$$F(p) = \int_a^b y(x)e^{-xp}dx, \qquad -\infty < a < b < \infty. \tag{1}$$

Then, for $y(x)$ continuous,

(α1) $F(p)$ is entire and of exponential type,

(α2) $|F(p)| \leqq M(e^{-a\sigma}+e^{-b\sigma})/|\sigma|$, $\sigma = \Re(p)$.†

Furthermore

$$y(x) = 1/2\pi i \int_{c-i\infty}^{c+i\infty} e^{xp}F(p)dp, \qquad c \text{ a real positive constant.} \tag{1.1}$$

For much of our work we shall be interested in the system‡

* Presented to the Society, November 28, 1936, under the title *A method for certain two-point boundary value problems.*

† Extensions involving the Riemann-Lebesgue lemma or differentiability properties of $y(x)$ are unnecessary for our present purpose.

‡ The differential equation terminology in this article is consistent with that of Ince, *Ordinary Differential Equations.* This book is designated by I in subsequent references.

$$(2) \qquad L(x, d/dx)y \equiv \sum_{i=0}^{n} a_i(x)d^i/dx^i y(x) = r(x),$$

$$(2.1) \qquad U_i(a, b) \equiv \sum_{j=0}^{n-1} \alpha_{ij}y^j(a) + \beta_{ij}y^j(b) = d_i, \qquad i = 1, \cdots, n,$$

where $a_i(x)$ is linear, α_{ij}, β_{ij}, d_i are constants and (2) is of rank 1.* In the sequel, bars over operators will denote adjoints and $P(\ )$ the bilinear concomitant. Furthermore, we assume: (A) $a_n(x) \neq 0$ for $a \leqq x \leqq b$, (B) $r(x)$ is piecewise continuous, (C) the rectangular matrix $\|\alpha_{ij}, \beta_{ij}\|$ is of rank n. We make a customary restriction: (D) to solutions continuous together with their first $n-1$ derivatives for $-\infty < a \leqq x \leqq b < \infty$. Restrictions (A), $\cdots$, (D) guarantee, also, piecewise continuity of the first derivative (that is, $y'(x)$) which is *sufficient here for the ordinary convergence of the right-hand integral in* (1.1) *to* $y(x)$. The relation $y(x) \doteq F(p)$ connotes the simultaneous validity of (1) and (1.1).†

The constant coefficient case already brings out the utility of relations (α). Here

$$(2.2) \qquad L(d/dx). \equiv \sum_{0}^{n} a_i d^i/dx^i., \quad L(d/dx). = L(-d/dx).,$$

$$(2.21) \qquad \int_a^b \{e^{-xp}L(d/dx)y(x) - y(x)\bar{L}(d/dx)e^{-xp}\}dx = P(y^i(x), p, e^{-xp})\Big|_a^b.$$

Therefore

$$(2.3) \qquad F(p)L(p) = R(p),$$

where

$$R(p) = \int_a^b e^{-xp}r(x)dx - P(y^i(x), p, e^{-xp})\Big|_a^b.$$

Let the J distinct roots $\{p_i\}$ of $L(p)$ be of multiplicities $\{m_i\}$; then $\sum_{i=1}^{J} m_i = n$. The root p_j must be a zero of $R(p)$ of order m_j at least, since $F(p)$ is entire. Thus the following n equations are satisfied:

* I, p. 427.

† For the notation and the facts used here concerning the Laplace-Mellin representation of the Heaviside calculus, cf. for instance D. G. Bourgin and R. J. Duffin, American Journal of Mathematics, vol. 59 (1937), p. 489. (Cf. G. Doetsch, *Theorie und Anwendung der Laplace Transformation*, Berlin, 1937, and the papers of D. V. Widder for further properties.)

$$(2.4)\qquad d^\sigma/dp^\sigma R(p)\Big|_{p=p_j} = 0, \qquad \sigma = 0, \cdots, m_j - 1; j = 1, \cdots, J,$$

with the convention $d^0/dp^0 R(p) = R(p)$.

The $2n$ equations, (2.1) and (2.4), are linear in the unknowns $y(a), \cdots, y^{(n-1)}(a), y(b), \cdots, y^{(n-1)}(b)$. If the coefficient and augmented matrices are of rank $2n-k$, the general solutions for these unknowns, in terms of $\alpha_{ij}, \beta_{ij}, d_i, a_i, p_j$ and $\int_a^b x^\sigma e^{-p_j x} y(x)dx$, involve k essential parameters.

The solution required is therefore

$$(2.5)\qquad y(x) \doteq [R(p)]/L(p) \ (\equiv F(p)).$$

The brackets denote here, and in (2.51), that the $y(a), \cdots, y^{n-1}(b)$ occurring in $R(p)$ are replaced by their determinate values (as found here from (2.1) and (2.4)). This method of solution is a special case of Method 1, described below.

We now take up in detail the case of linear coefficients. The analysis may be extended in an obvious manner to the case of polynomial coefficients for the differential equation of rank 1. We have

$$(3.1)\qquad \begin{aligned} L(x, d/dx)y(x) &\equiv L_0(d/dx)y(x) + xL_1(d/dx)y(x) \\ &\equiv \sum_{j=0}^{n} (a_j + xb_j)d^j/dx^j y(x) = r(x), \end{aligned}$$

with $b_n \neq 0$. Thus

$$L(x, d/dx). = L_0(-d/dx). + L_1(-d/dx)x.,$$

$$(3.2)\qquad \int_a^b \{e^{-xp}L(x, d/dx)y(x) - y(x)\bar{L}(x, d/dx)e^{-xp}\}dx = P(x, y^i(x), p, e^{-xp})\Big|_a^b,$$

that is,

$$(3.21)\qquad \int_a^b y(x)\bar{L}(x, d/dx)e^{-xp}dx = \int_a^b e^{-xp}r(x)dx - P(\ )\Big|_a^b \equiv R(p).$$

A straightforward reduction of the left side of (3.21) yields

$$(3.3)\qquad M(p, d/dp)F(p) \equiv L_0(p)F(p) - \frac{d}{dp}(L_1(p)F(p)) = R(p).$$

Let $G(p)$ satisfy

$$(3.4)\qquad \bar{M}(p, d/dp)G(p) \equiv \{L_0(p) + L_1(p)d/dp\}G(p) = 0,$$

that is,

$$G(p) = \exp\left[-\int^{p} L_0(p)/L_1(p)dp\right]. \tag{3.41}$$

Then

$$\int_{\gamma} \{G(p)M(p, d/dp)F(p) - F(p)\overline{M}(p, d/dp)G(p)\}dp = G(p)F(p)L_1(p)\Big|_{\gamma},$$

or

$$\int_{\gamma} G(p)R(p)dp = G(p)F(p)L_1(p)\Big|_{\gamma}, \tag{3.5}$$

where γ is to be chosen so that $F(p)G(p)L_1(p)\,|_{\gamma}$ vanishes.

Suppose the roots of $L_1(p)$ are $\{p_i\}$, $i=1, \cdots, l$, where p_i is of multiplicity $m_i>1$ for $i=1, \cdots, r$ and is a simple root for $r<i\leqq l$. We may write (3.41) in the form

$$G(p) = \exp\left[sp + \sum_{i=1}^{r}\sum_{j=1}^{m_i-1} B_{ij}/(p-p_i)^j\right](p-p_1)^{A_1}\cdots(p-p_l)^{A_l},$$

where $s=-a_n/b_n$. In view of (A) we may assume $s<a$ (though it may be necessary to multiply both sides of (3.1) by -1 to achieve this situation). The $\{A_k\}$ and $\{B_{ij}\}$ are constants, some or all of which may be 0 when $L_0(p)$ and $L_1(p)$ have roots in common.

The paths $\{\gamma'\}$ obtained by the usual application of the Laplace transformation to the homogeneous equation associated with (3.1) are the same as those obtained from its adjoint and satisfy*

$$e^{-xp}G(p)L_1(p)\Big|_{\gamma'} = 0. \tag{3.51}$$

So far as (3.5) is concerned, $F(p)$ need be considered solely in connection with approach to ∞ and indeed (cf. (α)) may be replaced by $e^{-a\sigma}$ (thus justifying termination of the path at $p=\infty$). (The factor $|\sigma|^{-1}$ is unimportant since $s\neq a, b$. Cf. (A).) Hence, on identifying x in (3.51) with a, it is plain that the often catalogued available paths† $\{\gamma'\}$ and $\{\gamma\}$ are the same when restriction is made to the case $|\Im(p)| < \infty$ for points on a path.

The function $R(p)$ of (3.21) may be written

$$(4)\quad e^{-xp}\sum_{j=0}^{n-1} y^{(j)}(x)\{L_0^{j+1}(p) + (x - d/dp)L_1^{j+1}(p)\}\Big|_a^b - \int_a^b e^{-xp}r(x)dx,$$

* I, p. 438.

† I, pp. 455, 457, 459.

where

$$p^j L_\alpha^j(p), \qquad \alpha = 0, 1, \tag{5}$$

represents the remainder of $L_\alpha(p)$ after the j terms of lowest degree are removed. From (3.5) we have

$$\begin{aligned} &y(b)F_{l0}(b) + \cdots + y^{(n-1)}(b)F_{l\,n-1}(b) - y(a)F_{l0}(a) - \cdots \\ &\qquad - y^{(n-1)}(a)F_{l\,n-1}(a) = \int_{\gamma_l} G(p) \int_a^b e^{-xp} r(x)\,dx\,dp, \end{aligned} \tag{6}$$

where

$$F_{lj}(z) = \int_\gamma e^{-zp} G(p) \{ L_0^{j+1}(p) + zL_1^{j+1}(p) - d/dp L_1^{j+1}(p) \}\,dp, \qquad l = 1, \cdots, n.$$

Relation (6) is central in our work. For applications we need

$$\left| \begin{matrix} \cdot & \cdot & \cdot \\ \cdot & F_{lj}(x) & \cdot \\ \cdot & \cdot & \cdot \end{matrix} \right| \neq 0, \qquad a \leqq x \leqq b. \tag{6.1}$$

We proceed now to establish relation (6.1). We introduce the notation (superscripts in parentheses indicate differentiation here and in the sequel)

$$V_l(x) = \int_{\gamma_l} e^{-xp} G(p)\,dp,$$

$$\omega(x) = \left\| \begin{matrix} V_1^{(n-1)}(x) & \cdots & V_1^{(k)}(x) & \cdots & V_1(x) \\ \cdot & \cdots & \cdot & \cdots & \cdot \\ V_n^{(n-1)}(x) & \cdots & V_n^{(k)}(x) & \cdots & V_n(x) \end{matrix} \right\|. \tag{6.2}$$

The determinant of this matrix is the Wronskian of $\{ V_l(x) \}$ and is denoted by $W(x)$. It may be shown directly that

$$\int_{\gamma_l} e^{-xp} p^i G(p)\,dp = (-1)^i d^i/dx^i V_l(x), \tag{6.3}$$

$$\bar{L}(x, d/dx) V_l(x) = 0, \tag{6.4}$$

$$(a_n + xb_n)(-1)^j V_l^{(j)}(x) = \sum_{\sigma=0}^{j-1} k_\sigma(x) F_{l\sigma}(x) + F_{lj}(x), \tag{6.5}$$

* The constants appearing are $y(a), \cdots, y^{n-1}(b)$ and not "undetermined" coefficients as in the usual methods. Thus, whenever, as in the one point problem, the value of $y(a)$ and the $n-1$ derivatives are given explicitly, the solution may be written down at once without the necessity of calculation of the values of the "undetermined" or arbitrary constants required in the non-Heaviside type procedures.

where $|k_q(x)| < \infty$, $q = 0, \cdots, n-1$, for $x > s$. From (6.5) we derive

$$\left| \begin{matrix} \vdots \\ \cdot \; F_{ij}(x) \; \cdot \\ \vdots \end{matrix} \right| = (-a_n - xb_n)^n W(x). \tag{6.6}$$

For large x, n contours $\{\gamma_i\}$ and solutions $\{V_i(x)\}$ of (6.5) always exist for which $W(x) \neq 0$.* Hence $W(x) \neq 0$ for $a \leqq x \leqq b$ by virtue of the Abel identity.† Since $s < a$, the verification of (6) is complete.

Method 1. Equations (2.1) and (6) provide a linear simultaneous system in the unknowns $y(a), \cdots, y^{(n-1)}(a), \cdots, y^{(n-1)}(b)$. By (6.1) and (6.6) the coefficient determinant may be written in the form

$$\Delta(a, b) = \left| \begin{matrix} & \vdots & \\ F_{rs}(a) & \cdot & F_{rs}(b) \\ & \vdots & \\ \cdots\cdots & \cdots & \cdots\cdots \\ & \vdots & \\ \alpha_{ij} & \cdot & \beta_{ij} \\ & \vdots & \end{matrix} \right| \tag{7}$$

$$\equiv \left| \begin{matrix} & \vdots & \\ -\omega(a) & \cdot & \omega(b) \\ & \vdots & \\ \cdots\cdots & \cdots & \cdots\cdots \\ & \vdots & \\ \alpha'_{ij} & \cdot & \beta'_{ij} \\ & \vdots & \end{matrix} \right| \{(a_n + ab_n)(a_n + bb_n)\}^n,$$

where (cf. (6.5))

$$\alpha'_{ij} = \alpha_{ij} + \sum_{\sigma=0}^{j-1} k_\sigma(a)\alpha_{i\sigma}, \quad \beta'_{ij} = \beta_{ij} + \sum_{\sigma=0}^{j-1} k_\sigma(b)\beta_{i\sigma}.$$

For purposes of exposition we make the hypothesis (E): $\Delta(a, b) \neq 0$. We may therefore solve for the unknowns mentioned and substitute their values in $R(p)$, thus obtaining $[R(p)]$.

In view of (3.21) and (3.5)

$$G(p)F(p)L_1(p) = C + \int^p G(p)[R(p)]dp. \tag{7.1}$$

The left side of (7.1) vanishes under our conditions at $p = \infty$. Hence, compare (2.5),

$$y(x) \doteq F(p) = \int_\infty^p G(p)[R(p)]dp/G(p)L_1(p). \tag{2.51}$$

The integration contour is a simple curve avoiding the singularities

* Cf. I, references, pp. 444, 445.

† I, p. 119.

$\{p_i\}$ of $G(p)$ and branch cuts.* However, no difficulty over other branches of $F(p)$ can arise since, by (α), $F(p)$ is single valued.

Method 1 may be viewed fairly as a generalization of the Heaviside procedure. We outline the argument. We restrict attention to the case that condition (A) is satisfied, that $r(x) \subset L(a, \infty)$ and is continuous for $x \geqq a$, that $\{p_i\}$ are distinct roots of $L(p)$ for (2) and of $L_1(p)$ for (3.1), and that $\Re(p_i) < 0 \leqq a$.

Under these restrictions it is easily shown that $|y(x)| < M$ for $x \geqq a$.† The function $y^{(i)}(x)$ appears in $P(\ )$ only as part of a product $e^{-xp}y^{(i)}(x)q_i(x)$, where $q_i(x)$ is a polynomial. Accordingly

$$L_{b\to\infty}P(y^i(b), p, e^{-bp}) = 0. \tag{7.2}$$

Since $y(a), \cdots, y^{(n-1)}(a)$ are known here and (7.2) justifies omission of $y(b), \cdots, y^{(n-1)}(b)$ for $b \to \infty$, we can write down $[R(p)]$ in (2.5) or (2.51) immediately without the mediation of the contours $\{\gamma_l\}$. (Evidently the Laplace integral in (1) exists for $\Re(p) > 0$ with b replaced by ∞, in virtue of the finiteness of $y(x)$; and, in fact, $y(x) \doteq F(p)$, since $y^{(1)}(x)$ is easily shown to be continuous for $x \geqq a$.) In the constant coefficient equation, *the result is identical with the Bromwich or Mellin integral formuation of the classic Heaviside solution, extended to nonzero data, for the one point problem.*

Method 2. An alternative solution of the two-point problem is available. On introducing the intermediate point X in (6) we may write

$$\sum_{k=0}^{n-1} y^{(k)}(x)F_{lk}(x)\Big|_a^X = \int_{\gamma_l}\int_a^X e^{-xp}r(x)dx, \qquad l = 1, \cdots, n. \tag{8}$$

Relations (6) and (8), together with (2.1), provide $3n$ equations in the $3n$ unknowns, $y(a), \cdots, y^{(n-1)}(a)$, $y(b), \cdots, y^{(n-1)}(b)$, $y(X), \cdots, y^{(n-1)}(X)$. The denominator determinant in the solution for $y(X)$ by Cramer's rule is reducible to $(-a_n - b_nX)^n \cdot W(X) \cdot \Delta(a, b)$. Under hypothesis (E) the unique solution $y(X)$ exists, satisfying the system (3.1) and (2.1).

Method 2 also yields a solution of the one point problem. Indeed if $\beta_{ij} = 0$ in (2.1), we solve for $y(X)$ from (2.1) and (8). The denominator $|F_{lj}(X)||\alpha_{ij}|$ and is nonzero for $X \geqq a$ by (C) and (6.1).

If $r(x) = 0$ and a parameter λ enters linearly in (2) or (3.1) or in (2.1), we have the characteristic function problem. The procedure for

* The particular choice of $\{p_{n_i}\}$ with $A_{n_i} > -1$ (where the direction of approach to p_{n_i} is determined by the signs of the associated B_{n_ij} if $B_{n_ij} \neq 0$), or of ∞ as the lower limit in (2.51) is of no consequence, since, for the case $A_{n_i} > -1$, the γ_l's, determined earlier, include the contours joining p_{n_i} and p_{n_j} or p_{n_i} and ∞ so that $\int_{p_{n_j},\infty}^{\infty, p_{n_i}} G(p)[R(p)]dp = 0$.

† I, p. 441, paragraph 18.12.

solution amounts to determining such values for λ that the algebraic systems (2.1) and (2.4) or (2.1) and (6) have a nonzero solution. This amounts to finding λ values for which $[R(p)]=0$. Methods 1 and 2 then apply. Since the analysis is straightforward enough we take up instead an instance of the sort of characteristic function problem arising in modern quantum theory, namely

$$L(x, d/dx)y \equiv xy^{(2)}(x) + (2 - xk^2)y(x) = 0, \tag{9}$$

$$y(x),\ y^{(1)}(x) \subset L(0, \infty);\ 3(k) = 0. \tag{9.1}^*$$

We note that (9) guarantees the continuity of $y(x)$, $y^{(1)}(x)$ and $y^{(2)}(x)$ for $x>a$. Equation (9.1) implies

$$L_{x\to 0}xy(x) = 0;\quad L_{x\to 0}xy^{(1)}(x) = 0;\ k^2 > 0. \tag{9.2}$$

Moreover the Laplace integrals of y, $y^{(1)}$, $xy^{(1)}$ certainly exist for $\mathcal{R}(p)>0$. Equation (9) is then the basis for the assertion that the Laplace integral of $xy^{(2)}$ exists for $\mathcal{R}(p)>0$. The analogues of relations (α) are (a) $F(p)$ is analytic for $\mathcal{R}(p)>0$; (b) $L_{\mathcal{R}(p)\to\infty}pF(p)=0$. Now

$$\int_{\bar{x}}^{\infty}\left\{e^{-xp}L(x, d/dx)y - y\bar{L}(x, d/dx)e^{-xp}\right\}dx = \left[e^{-xp}(y(xp - 1) + xy')\right]\Big|_{\bar{x}}^{\infty}. \tag{9.3}$$

Since the integrand is the algebraic sum of functions belonging to $L(0, \infty)$, the right-hand side of (9.3) is absolutely continuous in $\bar{x}$ for $0\leqq\bar{x}\leqq 1$, and the limit exists for $\bar{x}\to 0$, that is,

$$(p^2 - k^2)F^{(1)}(p) + 2(p - 1)F(p) = y(0+). \tag{9.4}$$

According to (7.1) we have then

$$F(p) = \frac{(p - k)^{-1+1/k}}{(p + k)^{1+1/k}}\left[C + y(0+)\int_{p_0}^{p}\left(\frac{p + k}{p - k}\right)^{1/k}dp\right], \tag{9.5}$$

where C and p_0, $p_0\neq\pm k$, are arbitrary finite constants. Since $F(p)$ is unaffected when k is replaced by $-k$, we may assume $k>0$. Clearly

$$L_{\mathcal{R}(p)\to\infty}\left|\int_{p_0}^{p}\left(\frac{z + k}{z - k}\right)^{1/k}dz\right| = O(|p|). \tag{9.6}$$

The contour avoids a possible branch cut from $-k$ to k. By (9.6)

$$L_{p\to\infty}|pF(p)| = |y(0+)O(1)|. \tag{9.7}$$

By (a), $F(p)$ is single valued. Hence, by (b), $y(0+)$ vanishes.

* The radial part of the hydrogen wave equation for an S state.

According to (a), $p=k$ cannot be a branch point. Therefore

$$k = 1/n, \qquad n = 1, 2, \cdots. \tag{9.8}$$

These are the characteristic values. The associated characteristic functions are given by

$$y_n(x) \doteq C'\left[\frac{(1-np)^{n-1}}{(1+np)^{n+1}}\right] \doteq C''(1 - nd/dx)(x^n e^{-x/n}). \tag{9.9}$$

The extension of Method 2 to partial differential equations is striking, though not directly obvious. We treat the equation

$$L\left(x, \frac{\partial}{\partial x}, \frac{\partial}{\partial t}\right) y(x, t) \equiv (a + x)y_{xx}(x, t) + by_{tt}(x, t) = r(x, t), \qquad a, b > 0, \tag{10}$$

with data assigned on the boundaries of the rectangle

$$0 \leqq x \leqq S, \quad 0 \leqq t \leqq T. \tag{10.1}$$

Our procedure yields the formal expansion

$$y(x, t) = \sum_1^\infty Y(n; x) \sin n\pi t/T; \tag{10.2}$$

$$y(x, -t) = y(x, 2T - t) = -y(x, t). \tag{10.3}$$

If (F): the boundary values $y(x, 0)$, $y(x, T)$ vanish, the individual terms in the series of (10.2) satisfy the differential equation. In the contrary case this need no longer be true. The latter situation is easily clarified. We discuss the case that the series of (10.2) is a Fourier series for fixed x. Because of the discontinuity of the extended function $y(x, t)$ (10.3) for $t=0$, T, the coefficients $Y(n; x)$ are at best $O(n^{-1})$.* Accordingly, the series of (10.2) cannot be differentiated term by term when (F) is not satisfied.

We deal first with a subrectangle contained between the ordinates $x=0$, $x=X \leqq S$. We write

$$F(X; p, u) = \int_0^T \int_0^X e^{-xp+iut} y(x, t)dxdt. \tag{1'}$$

The operator adjoint to $L(x, \cdots)$ is $\overline{L}(x, \partial/\partial x, \partial/\partial t) \equiv \partial^2(a+x)/\partial x^2. +b\partial^2/\partial t^2..$ There results from Green's theorem (compare (3.3))

* Courant-Hilbert, *Methoden der Mathematischen Physik*, 1st edition, p. 53. A similar situation is encountered in the usual Dirichlet problem for the rectangle when the solution is expressed as a double Fourier sine series.

$$- M(u, p, \partial/\partial p)F(X; p, u) + \int_0^T \int_0^X e^{-xp+iut} r(x, t)dxdt$$

$$(3.2') \qquad = \int_0^T \{(a + x)(y_x + py) - y\} e^{-xp+iut} dt \Big|_{x=0}^{x=X}$$

$$+ b \int_a^X \{(y_t - iuy)e^{-px+iut}\} dx \Big|_{t=0}^{t=T};$$

$$M(u, p, \partial/\partial p). = \partial/\partial p p^2. - (bu^2 - ap^2).$$

Clearly $G(p, u) = \exp -(ap + bu^2/p)$ satisfies

$$\overline{M}(u, p, \partial/\partial p)G(p, u) = (- p^2 \partial/\partial p - (bu^2 - ap^2))G(p, u) = 0;$$

$$(3.51') \qquad - \int_\gamma G(p, u)M(u, p, \partial/\partial p)F(X; p, u)dp = p^2 F(X; p, u)G(p, u)\Big|_\gamma.$$

The conditions (α) are evidently applicable for finite u. The right-hand expression in (3.5′) vanishes for (a) γ_1: the positive real p axis, (b) γ_2: a circle in the p plane, with center at the origin and unit radius, described in a counter clockwise sense.

We take the special values $u = n\pi/T$ with n an integer and use the imaginary part* of (3.5′) where $M(u, p, \partial/\partial p)F(X; p, u)$ is replaced by its value according to (3.2′). The inner reason for the success of our procedure is that the y_t term (cf. (3.2′)) drops out. We have

$$Y_x(n, X)V^\sigma(n, X) - Y_x(n, 0)V^\sigma.(n, 0) - Y(n, X)V_x^\sigma(n, X) = Z^\sigma(n, X), \qquad \sigma = 1, 2,$$

$$Y(n, X) = \int_0^T y(X, t) \sin n\pi t/T dt,$$

$$V^\sigma(n, X) = (2\pi i)^{1-\sigma}(a + X) \int_{\gamma_\sigma} \exp\left[- (a + X)p - bn^2\pi^2/pT^2\right]dp,$$

$$Z^\sigma(n, X) = - \dot{Y}(n, 0)V_x^\sigma(n, 0)$$

$$- bn\pi/T \int_0^X \frac{V^\sigma(n, x)}{a + x} (y(x, 0) + (- 1)^{n-1} y(x, T))dx$$

$$+ \int_0^T \int_0^X \frac{V^\sigma(n, x)}{a + x} \sin \frac{n\pi t}{T} r(x, t)dxdt.$$

The application of the preceding analysis to the complete rectangle of (10.1) yields equations, to be denoted as (1″) and (6″), in which X is replaced by S. We may solve (6′) and (6″) formally, by

* The cosine integrals arising from the real part solve the Neumann problem.

Cramer's rule, for $Y(n, X)$ in terms of integrals of the boundary values and of $r(x, t)$, namely

$$
(10.4)\qquad
\begin{aligned}
Y(n, X) = {} & \begin{vmatrix}
Z^1(n, X) & V^1(n, X) & 0 & V^1(n, 0) \\
Z^2(n, X) & V^2(n, X) & 0 & V^2(n, 0) \\
Z^1(n, S) + Y(n, S)V_x^1(n, S) & 0 & V^1(n, S) & V^1(n, 0) \\
Z^2(n, S) + Y(n, S)V_x^2(n, S) & 0 & V^2(n, S) & V^2(n, 0)
\end{vmatrix} \\
& \div \begin{vmatrix} V^1(n, X) & V_x^1(n, X) \\ V^2(n, X) & V_x^2(n, X) \end{vmatrix} \begin{vmatrix} V^1(n, 0) & V^1(n, S) \\ V^2(n, 0) & V^2(n, S) \end{vmatrix}.
\end{aligned}
$$

It may be shown that

$$
(10.5)^*\qquad
\begin{aligned}
V^1(n, X) &= (2n\pi/T)(a + X)^{1/2}K_{-1}((2n\pi/T)(a + X)^{1/2}), \\
V^2(n, X) &= -(2n\pi/T)(a + X)^{1/2}I_1((2n\pi/T)(a + X)^{1/2}).
\end{aligned}
$$

The first denominator determinant in (10.4) is the Wronskian of $V^1(n, X)$ and $V^2(n, X)$ and may be shown to be a nonzero constant. Either from (10.5) or the integral representation, it may be demonstrated that $V^1(n, X)/V^2(n, X)$ is monotone increasing for $X \geqq 0$, whence it follows that the second denominator determinant in (10.4) cannot be 0. Thus the solution for $y(n, X)$ is not illusory.

The expansion in (10.2) with coefficients given by (10.4) is an interesting extension of the generalized Schlomilch series, heretofore not treated in the literature.† In case (F), if the data go to 0 at $t = 0$ or T and are, for instance, of class C^2 and $r(x, t) = 0$, it is easy to establish the convergence of the series in (10.2) to a solution. However we defer the investigation of the properties of such expansions.‡

UNIVERSITY OF ILLINOIS AND
PURDUE UNIVERSITY

* G. N. Watson, *Theory of Bessel Functions*, pp. 80, 181, 183.

† Watson, ibid., p. 622.

‡ The introduction of an accessory line $x = X$ and subsequent elimination of derivatives by taking two regions is central in all applications of the second method to the hyperbolic and parabolic equations also. For mixed problems and a finite range on either variable the corresponding parameter is given discrete values; for the Cauchy problem both u and p are continuous. Method 1 is also available.

NOTE ON CERTAIN LAGRANGE INTERPOLATION POLYNOMIALS*

M. S. WEBSTER

Let

$$l_k(x) \equiv l_k^{(n)}(x) \equiv \frac{\phi_n(x)}{\phi_n'(x_k)(x - x_k)}, \qquad l_k(x_k) = 1,$$

where $\phi_n(x) \equiv (x-x_1)(x-x_2)\cdots(x-x_n)$; $x_k \equiv x_{k,n} \equiv \cos\theta_k \equiv \cos\theta_{k,n}$; $0<\theta_1<\theta_2<\cdots<\theta_n<\pi$; $-1<x_k<1$, $(k=1, 2, \cdots, n)$. The polynomials $l_k(x)$ of degree $n-1$ are the fundamental polynomials of Lagrange interpolation. In this note we suppose that $\phi_n(x) \equiv \phi_n(x; \alpha, \beta)$ is the Jacobi polynomial which satisfies the following differential equation:

$$(1 - x^2)\phi_n''(x) + [\alpha - \beta - (\alpha + \beta)x]\phi_n'(x) + n(n + \alpha + \beta - 1)\phi_n(x) = 0, \qquad n = 1, 2, \cdots,$$

where α, β are positive parameters.

We develop certain bounds and limiting relations for $l_k(x)$ for special values of α, β, obtaining results similar to those of Erdös, Grünwald, and Lengyel.

It is known that ([1], pp. 17, 31, 33, 35, 62):

$$\phi_n(-x; \beta, \alpha) = (-1)^n\phi_n(x; \alpha, \beta), \quad \phi_n'(x; \alpha, \beta) = n\phi_{n-1}(x; \alpha + 1, \beta + 1),$$

$$x_{k,n}(\beta, \alpha) = -x_{n-k+1,n}(\alpha, \beta), \quad l_k^{(n)}(x; \beta, \alpha) = l_{n-k+1}^{(n)}(-x; \alpha, \beta),$$

$$\int_{-1}^{1} (1 + x)^{\alpha-1}(1 - x)^{\beta-1}\phi_m(x)\phi_n(x)dx = 0, \qquad m \neq n,$$

$$\phi_n(x) = \frac{(-2)^n\Gamma(n + \alpha + \beta - 1)\Gamma(n + 1/2)}{(\pi)^{1/2}\Gamma(2n + \alpha + \beta - 1)} (\sin\lambda)^{1/2-\alpha}(\cos\lambda)^{1/2-\beta} \cdot \left\{\cos\left[(2n + \alpha + \beta - 1)\lambda - \frac{\pi}{4}(2\alpha - 1)\right] + O\left(\frac{1}{n}\right)\right\},$$

$$\sin^2\lambda = \frac{1 + x}{2}; \quad -1 + \epsilon \leqq x \leqq 1 - \epsilon, \ \epsilon > 0.$$

If $\alpha = \beta = 1/2$, $\phi_n(x) = (1/2^{n-1})\cos n\theta$, $(x = \cos\theta)$, which is the Tschebycheff polynomial of degree n.

The following lemma due to M. Riesz [2] will be useful:

* Presented to the Society, September 7, 1939.

LEMMA. *A trigonometric polynomial of degree n assumes the maximum of its absolute value at a point whose distance from any of the roots of the trigonometric polynomial is not less than $\pi/2n$.*

Let $\alpha=\beta=3/2$. Then, we have ([1], p. 17)

$$\phi_n(x)=\frac{\sin(n+1)\theta}{2^n\sin\theta},\qquad x=\cos\theta,$$

$$l_k(x)=\frac{(-1)^{k+1}\sin^2\theta_k\sin(n+1)\theta}{(n+1)\sin\theta(\cos\theta-\cos\theta_k)},\qquad \theta_k=\frac{k\pi}{n+1}.$$

Since $\theta_{k+1}-\theta_k=\pi/(n+1)$, the lemma shows that the maximum of $|l_k(\cos\theta)|$ occurs between θ_{k-1} and θ_{k+1} or at $x=\pm1$, $(k=2, 3, \cdots, n-1)$. Also, $|l_1(\cos\theta)|$ and $|l_n(\cos\theta)|$ attain their maxima at $\theta=0$ and $\theta=\pi$, respectively. We find that

$$l_1^{(1)}(x)=1,\qquad |l_1^{(2)}(x)|=|x+1/2|\leqq 3/2,$$
$$|l_2^{(2)}(x)|=|-x+1/2|\leqq 3/2,$$
$$l_1^{(n)}(1)=l_n^{(n)}(-1)=2\cos^2\pi/[2(n+1)],$$
$$\lim_{n\to\infty}l_1^{(n)}(1)=\lim_{n\to\infty}l_n^{(n)}(-1)=2.$$

In the following we may assume $n\geqq 3$, $2\leqq k\leqq n-1$ and $x_k\geqq 0$. We have

$$\phi_n'(x)=-\frac{(n+1)\sin\theta\cos(n+1)\theta-\cos\theta\sin(n+1)\theta}{2^n\sin^3\theta},$$

$$\phi_n''(x)=-\frac{n(n+2)\sin^2\theta\sin(n+1)\theta+3(n+1)\sin\theta\cos\theta\cos(n+1)\theta}{2^n\sin^5\theta}$$
$$-\frac{3\cos^2\theta\sin(n+1)\theta}{2^n\sin^5\theta}.$$

Since $\phi_n''(x)$ has opposite signs at x_k and $\cos[(2k-1)\pi/(n+1)]$, we must have $\phi_n''(\bar{x})=0$ where $\bar{x}\equiv\cos\bar{\theta}\equiv\cos(k-\epsilon)\pi/(n+1)$, where $0\leqq\epsilon<1/2$.

It follows from the differential equation that $3\bar{x}\phi_n'(\bar{x})=n(n+2)\cdot\phi_n(\bar{x})$. Since $|\phi_n(\bar{x})|=|(\bar{x}-x_k)\phi_n'(\gamma)|\geqq|(\bar{x}-x_k)\phi_n'(x_k)|$, $(x_k<\gamma<\bar{x})$, we have

$$0\leqq\bar{x}-x_k\leqq\frac{3\bar{x}}{n(n+2)}\left|\frac{\phi_n'(\bar{x})}{\phi_n'(x_k)}\right|.$$

If

$$\max_{x_{k+1}\leq x\leq x_{k-1}} |l_k^{(n)}(x)| = |l_k^{(n)}(\mu)|,$$

we have

$$\max_{x_{k+1}\leq x\leq x_{k-1}} \left|\frac{\phi_n(x)}{\phi_n'(x_k)(x-x_k)}\right| = \left|\frac{\phi_n'(\delta)}{\phi_n'(x_k)}\right| \leqq \left|\frac{\phi_n'(\bar{x})}{\phi_n'(x_k)}\right| \equiv u,$$

$$x_k < \delta < \mu < x_{k-1}.$$

Now $(n+1)\tan\bar{\theta} > 3\pi/2$ and

$$\left|\frac{\sin(k-\epsilon)\pi}{(n+1)\tan\bar{\theta}}\right| < \frac{2}{3\pi}$$

so that

$$\begin{aligned} u &= \frac{1-x_k^2}{1-\bar{x}^2}\left[\cos(k-\epsilon)\pi - \frac{\sin(k-\epsilon)\pi}{(n+1)\tan\bar{\theta}}\right] \\ &< c\left[1+\frac{(\bar{x}+x_k)(\bar{x}-x_k)}{1-\bar{x}^2}\right] < c\left[1+\frac{6\bar{x}^2}{n(n+2)(1-\bar{x}^2)}\,u\right] \\ &< \frac{c}{1-\dfrac{6c\bar{x}^2}{n(n+2)(1-\bar{x}^2)}} = \frac{c}{1-\dfrac{6c}{n(n+2)\tan^2\bar{\theta}}} \\ &< 1.87, \end{aligned}$$

where $c = 1+2/3\pi$. Since $|l_k(1)| = 1+x_k < 2$, $|l_k(-1)| = 1-x_k < 2$, we have proved [3] the following theorem.

THEOREM 1. *If* $\alpha=\beta=3/2$ *so that* $\phi_n(x)=\sin(n+1)\theta/2^n\sin\theta$, $(x=\cos\theta)$, *then* $|l_k^{(n)}(x)|<2$, $(-1\leqq x\leqq 1)$, *and* $\lim_{n\to\infty} l_1^{(n)}(1) = \lim_{n\to\infty} l_n^{(n)}(-1) = 2$. *Also,* $\lim_{n\to\infty}|l_k^{(n)}(1)| = 2$ *if, and only if,* $x_{k,n}\to 1$ *as* $n\to\infty$, *and* $\lim_{n\to\infty}[\max_{-1\leq x\leq 1}|l_k^{(n)}(x)|] = 2$ *if, and only if,* $x_{k,n}\to\pm 1$ *as* $n\to\infty$.

Using the same method we may prove the following similar theorems.

THEOREM 2. *If* $\alpha=1/2$, $\beta=3/2$ *so that*

$$\phi_n(x) = \frac{\sin(2n+1)(\theta/2)}{2^n\sin(\theta/2)}, \qquad x=\cos\theta,$$

then $|l_k^{(n)}(x)|<2$, $(-1\leqq x\leqq 1)$, *and* $\lim_{n\to\infty} l_1^{(n)}(1)=2$, $\lim_{n\to\infty} l_n^{(n)}(-1) = 4/\pi$.

THEOREM 3. *If* $\alpha=3/2, \beta=1/2$ *so that*

$$\phi_n(x) = \frac{\cos(2n+1)(\theta/2)}{2^n \cos(\theta/2)}, \qquad x = \cos\theta,$$

then $|l_k^{(n)}(x)| < 2$, $(-1 \leqq x \leqq 1)$, $\lim_{n\to\infty} l_1^{(n)}(1) = 4/\pi$, $\lim_{n\to\infty} l_n^{(n)}(-1) = 2$.

It should be noted that the upper bound 2 in these three theorems is the least upper bound for which the inequalities remain valid.

Let ϵ be an arbitrary, fixed positive number less than 1. We suppose that x_k is restricted to an interval I where $-1+\epsilon \leqq x_k \leqq 1-\epsilon$. Using the asymptotic formula for $\phi_n(x)$ we easily obtain the following theorem.

THEOREM 4. *For all* x *and* x_k *in* I *and* $|x-x_k| \geqq \epsilon' > 0$, $|l_k^{(n)}(x)| = O(1/n)$ *as* $n\to\infty$.

Hence, if $|l_k^{(n)}(x)|$ attains its maximum in I at $x=\mu\equiv\mu_{k,n}$, we must have $\mu-x_k\to 0$ as $n\to\infty$. At $x=\mu$, we have $(\mu-x_k)\phi_n'(\mu)-\phi_n(\mu)=0$ so that

$$\max |l_k^{(n)}(x)| = \left|\frac{\phi_n'(\mu)}{\phi_n'(x_k)}\right| = \left|\frac{n\phi_{n-1}(\mu;\alpha+1,\beta+1)}{n\phi_{n-1}(x_k;\alpha+1,\beta+1)}\right| \to 1$$

as $n\to\infty$, provided x and x_k are in I. This proves the following theorem.

THEOREM 5. *For all* x_k *such that* $-1+\epsilon \leqq x_k \leqq 1-\epsilon$, $(\epsilon>0)$, $\max_{-1+\epsilon\leq x\leq 1-\epsilon}|l_k^{(n)}(x)| \to 1$ *as* $n\to\infty$.

There is a close connection between Theorems 4 and 5 and the results of Erdös and Lengyel [4].

REFERENCES

1. J. Shohat, *Théorie générale des polynômes orthogonaux de Tchebichef*, Mémorial des Sciences Mathématiques, vol. 66, Paris, 1934.

2. M. Riesz, *Eine trigonometrische Interpolationsformel und einige Ungleichungen für Polynome*, Jahresbericht der deutschen Mathematiker-Vereinigung, vol. 23 (1914), pp. 354–368.

3. A. Schaeffer obtained a somewhat similar proof (unpublished) for Tschebycheff polynomials. Also, see Erdös and Grünwald, *Note on an elementary problem of interpolation*, this Bulletin, vol. 44 (1938), pp. 515–518.

4. P. Erdös and B. Lengyel, *On fundamental functions of Lagrangean interpolation*, this Bulletin, vol. 44 (1938), pp. 828–834.

PURDUE UNIVERSITY

ON FINITE ABELIAN p-GROUPS*

DOUGLAS DERRY

Let p denote a prime number which is fixed throughout the course of the paper. As we shall deal only with abelian groups whose order is some power of p, the term group will be used exclusively to designate a group of this nature.

DEFINITION. *A subgroup $\mathfrak{H}$ of a group $\mathfrak{G}$ is said to be reciprocally cyclic if the factor group $\mathfrak{G}/\mathfrak{H}$ is cyclic.*

The letters k, m, n, r, x will be used to denote nonnegative integers and the symbol $(S_1, \cdots, S_n)$ to denote the group generated by the elements $S_1, \cdots, S_n$.

DEFINITION. *A set of elements $S_1, \cdots, S_n$ of the group $\mathfrak{G}$ is said to define a composition series for $\mathfrak{G}$ if the group series*

$$(S_1, \cdots, S_n), (S_1, \cdots, S_{n-1}), \cdots, (S_1, S_2), (S_1), (E)$$

is a composition series for the group $\mathfrak{G}$.

DEFINITION. *A set of elements $S_1, S_2, \cdots, S_n$ of the group $\mathfrak{G}$ is said to form a coverage system of the group $\mathfrak{G}$ if for any reciprocally cyclic subgroup $\mathfrak{H}$ of $\mathfrak{G}$ we may select elements $S_{m_1}, \cdots, S_{m_k}$ so that the group series*

$$(S_{m_1}, \cdots, S_{m_k}, \mathfrak{H}), (S_{m_1}, \cdots, S_{m_{k-1}}, \mathfrak{H}), \cdots, (S_{m_1}, \mathfrak{H}), \mathfrak{H}$$

is a composition series from $\mathfrak{G}$ to $\mathfrak{H}$.

A set of elements $S_1, \cdots, S_n$ which defines a composition series for a group $\mathfrak{G}$ also forms a coverage system for $\mathfrak{G}$. For let

$$(S_1, \cdots, S_n), (S_1, \cdots, S_{n-1}), \cdots, (S_1, S_2), (S_1), (E)$$

be the composition series of $\mathfrak{G}$ resulting from the element system $S_1, \cdots, S_n$ and $\mathfrak{H}$ any reciprocally cyclic subgroup of $\mathfrak{G}$; then the series

$$(S_1, \cdots, S_n, \mathfrak{H}), (S_1, \cdots, S_{n-1}, \mathfrak{H}), \cdots, (S_1, \mathfrak{H}), \mathfrak{H}$$

is a composition series from $\mathfrak{G}$ to $\mathfrak{H}$ with possible repetitions. These may be removed by starting with S_1 and striking out successively from left to right each member of the set $S_1, \cdots, S_n$ which is unnecessary for the generation of the groups of the above series. The

* Presented to the Society, April 14, 1939.

elements remaining define a composition series from $\mathfrak{G}$ to $\mathfrak{H}$ which shows the set to be a coverage system for the group $\mathfrak{G}$.

This investigation deals with the converse of the above problem which may be stated as follows: $\mathfrak{G}$ is a group with composition series of length $n+1$ and with coverage system $S_1, \cdots, S_n$. Does it follow that this coverage system (after a possible rearrangement of order of the terms) defines a composition series for $\mathfrak{G}$? For groups of rank 1, that is, for cyclic groups the answer is clearly in the affirmative. The present paper shows the result to be true for groups of rank 2. The problem arises in the theory of numbers but is closely connected with ideas involved in the Jordan-Hölder theorem.

THEOREM A. *$\mathfrak{H}$ is any subgroup of the group $\mathfrak{G}$. The elements $S_1, \cdots, S_n$ form a coverage system of the group $\mathfrak{G}$. Then the restclasses defined by these elements in the factor group $\mathfrak{G}/\mathfrak{H}$ define a coverage system for this group.*

PROOF. Let $\overline{\mathfrak{F}}$ be a reciprocally cyclic subgroup of the factor group $\mathfrak{G}/\mathfrak{H}$. If $\mathfrak{F}$ be the maximum subgroup of the group $\mathfrak{G}$ which is built into $\overline{\mathfrak{F}}$ in the homomorphism $\mathfrak{G}\sim\mathfrak{G}/\mathfrak{H}$, it follows from the second isomorphism theorem that $\mathfrak{G}/\mathfrak{F}\cong\mathfrak{G}/\mathfrak{H}/\overline{\mathfrak{F}}$. This latter group is by hypothesis cyclic. Consequently $\mathfrak{F}$ is a reciprocally cyclic subgroup of $\mathfrak{G}$. Then by the definition of the coverage system, elements $S_{m_1}, S_{m_2}, \cdots, S_{m_k}$ exist so that the series

$$(S_{m_1}, S_{m_2}, \cdots, S_{m_k}, \mathfrak{F}), \cdots, (S_{m_1}, \mathfrak{F}), \mathfrak{F}$$

is a composition series from $\mathfrak{G}$ to $\mathfrak{F}$.

Let $\overline{S}_{m_1}, \overline{S}_{m_2}, \cdots, \overline{S}_{m_k}$ denote the restclasses of $\mathfrak{G}/\mathfrak{H}$ defined by the elements $S_{m_1}, S_{m_2}, \cdots, S_{m_k}$, respectively. Now in a homomorphism a composition series is built into a composition series with possible repetitions. Accordingly, as in the homorphism $\mathfrak{G}\sim\mathfrak{G}/\mathfrak{H}$, the series

$$(S_{m_1}, S_{m_2}, \cdots, S_{m_k}, \mathfrak{F}), \cdots, (S_{m_1}, \mathfrak{F}), \mathfrak{F}$$

is built into the series

$$(\overline{S}_{m_1}, \overline{S}_{m_2}, \cdots, \overline{S}_{m_k}, \overline{\mathfrak{F}}), \cdots, (\overline{S}_{m_1}, \overline{\mathfrak{F}}), \overline{\mathfrak{F}};$$

this second series is a composition series from $\mathfrak{G}/\mathfrak{H}$ to $\overline{\mathfrak{F}}$ with possible repetitions. Starting with the term $\overline{S}_{m_1}$ and proceeding from left to right, we successively remove every term of the set $\overline{S}_{m_1}, \overline{S}_{m_2}, \cdots, \overline{S}_{m_k}$ which is unnecessary for the generation of the groups of this series. After the repetitions have been thus removed, the series is a composition series from $\mathfrak{G}/\mathfrak{H}$ to $\overline{\mathfrak{F}}$ generated by the terms remaining of the set $\overline{S}_{m_1}, \overline{S}_{m_2}, \cdots, \overline{S}_{m_k}$. Thus the theorem is proved.

THEOREM B. *$\mathfrak{G}$ is a group for which the set $S_1, \cdots, S_n$ defines* (1) *a coverage system for $\mathfrak{G}$,* (2) *a composition series for $\mathfrak{G}$. S is an element of order p for which either $S=S_k$, or $S_k=SS_m^x$, $(p\nmid x,\ k>m)$. Then the restclasses defined by $S_1, \cdots, S_{k-1}, S_{k+1}, \cdots, S_n$ define* (1) *a coverage system for $\mathfrak{G}/(S)$,* (2) *a composition series for $\mathfrak{G}/(S)$.*

PROOF. Let us consider (1) first. By Theorem A the restclasses defined by $S_1, \cdots, S_n$ in the group $\mathfrak{G}/(S)$ form a coverage system. If $S=S_k$, the element S_k represents the unit restclass and so may be omitted from the coverage system. As for the second possibility, $S_k=SS_m^x$, we note that, as $p\nmid x$, S_m and S_m^x are equivalent as group generators, that is, we may always use S_m as a generator in place of S_m^x. In the factor group $\mathfrak{G}/(S)$, S_k and S_m^x represent the same restclass and consequently S_k may be omitted from the coverage system. Thus (1) is true. To prove (2) we use the above relations to show that the restclasses defined by $S_1, \cdots, S_{k-1}, S_{k+1}, \cdots, S_n$ define a composition series of $\mathfrak{G}/(S)$ with possible repetitions. As the length of the composition series of $\mathfrak{G}/(S)$ is n there can be no repetitions and so the result of the theorem is established.

The hypothesis of Theorems C and D is as follows. $S_1, \cdots, S_n$ is a system of elements, with S_1 the only element of order p, which defines a composition series for the noncyclic group $\mathfrak{G}$. The number a is the least subscript so that the group generated by the elements $S_1, \cdots, S_a$ is not cyclic; X is any element for which $X^p=E$ and $(S_{a-1}, X)=(S_{a-1}, S_a)$. Immediately from these relations follows

$$S_a = S_r^{x_1}X^{x_2}, \qquad p\nmid x_1,\ p\nmid x_2,\ 1<r<a. \tag{1}$$

Let $T_1=S_{m_1}$, $T_2=S_{m_2}, \cdots, T_k=S_{m_k}$ be a greatest possible subset chosen from $S_2, \cdots, S_n$ so that $X, T_1, \cdots, T_k$ defines a composition series for the subgroup $\mathfrak{G}_X=(X, T_1, \cdots, T_k)$. If the subset is void, $\mathfrak{G}_X=(X)$.

THEOREM C. *If Y satisfy the same conditions as X with $(X)\neq(Y)$ and $\mathfrak{G}_Y$ be defined analogously to $\mathfrak{G}_X$, then $\mathfrak{G}_X \cap \mathfrak{G}_Y=(E)$.*

PROOF. We first show

$$S_1 \notin \mathfrak{G}_X. \tag{2}$$

If this were false and $S_1 \in \mathfrak{G}_X$, then $S_2 \in \mathfrak{G}_X$, for otherwise from $(S_2, S_1):(S_1)=p$ would follow $(\mathfrak{G}_X, S_2):\mathfrak{G}_X=p$ and $X, T_1, T_2, \cdots, T_k$ would not be a greatest possible subset. We could in turn argue similarly $S_2 \in \mathfrak{G}_X$ and so for all elements $S_2, \cdots, S_n$. Hence $\mathfrak{G}_X=\mathfrak{G}$ and thus $X, T_1, \cdots, T_k$ would define a composition series for $\mathfrak{G}$. Then,

from the Jordan-Hölder theorem, $k=n-1$ and consequently $T_1, \cdots, T_k$ would be the complete set $S_2, \cdots, S_n$. Therefore, for S_r defined by (1), elements T_{m_1}, T_{m_2} would exist so that $T_{m_1}=S_r$, $T_{m_2}=S_a$. Now one of T_{m_1}, T_{m_2}, say T_{m_1}, must precede the other. Using (1) from Theorem B, it would then follow that $X, T_1, \cdots, T_{m_2-1}$, $T_{m_2+1}, \cdots, T_k$ would define a composition series for $\mathfrak{G}/(X)$ which is impossible as X is within the unit restclass. Therefore (2) must hold.

From the definition of a and the fact that $S_a^p \neq E$ follows $S_1 \epsilon (S_x)$, $(1 \leqq x \leqq a)$. Therefore $S_x \epsilon \mathfrak{G}_X$, $(1 \leqq x \leqq a)$, has as a consequence $S_1 \epsilon \mathfrak{G}_X$. This being false, we have

$$S_x \notin \mathfrak{G}_X, \qquad 1 \leqq x \leqq a. \tag{3}$$

LEMMA. *For* $S \epsilon (X, T_1, \cdots, T_{k_1})$, $(k_1 \leqq k)$, *a representation*

$$S = X^{r_0} S_{n_1}^{r_1} \cdots S_{n_{k_2}}^{r_{k_2}}, \qquad p \nmid r_1, p \nmid r_2, \cdots, p \nmid r_{k_2}; n_1 < n_2 < \cdots < n_{k_2},$$

exists where $S_{n_1}, \cdots, S_{n_{k_2}}$ *are members of the set* $T_1, \cdots, T_{k_1}$.

PROOF OF LEMMA. If $S \epsilon (X, T_1, \cdots, T_{k_1})$, integers $x_0, \cdots, x_{k_1}$, exist for which $S = X^{x_0} T_1^{x_1} \cdots T_{k_1}^{x_{k_1}}$. Let x_{k_2} be the first of these integers from the right for which $p \nmid x_{k_2}$, $x_{k_2} \neq 0$. Now $k_2 > 0$, otherwise there is nothing to prove. As $(X, T_1, \cdots, T_{k_2}) : (X, T_1, \cdots, T_{k_2-1}) = p$, $T_{k_2}^p \epsilon (X, T_1, \cdots, T_{k_2-1})$; X may be regarded as T_0. Therefore an expression for $T_{k_2}^{x_{k_2}}$ exists in terms of $X, T_1, \cdots, T_{k_2-1}$. Substituting this expression for $T_{k_2}^{x_{k_2}}$ into the above representation for S, we obtain a new representation for S in which the exponent of T_{k_2} is zero and the exponents of $T_{k_2+1}, \cdots, T_{k_1}$ are the same as before. By successively repeating this process we ultimately arrive at an expression for S in which the exponents of $T_1, \cdots, T_{k_1}$ are either zero or prime to p. Omitting the factors with zero exponents, we now substitute for those remaining their values in the set $S_1, \cdots, S_n$, thus obtaining after a possible rearrangement of order the representation of S in the lemma.

Now let us assume the theorem false, that is, the meet $\mathfrak{G}_X \cap \mathfrak{G}_Y$ contains a non-unit element. From (1) and the analogous expression in Y we obtain $Y = S_1^{z_1} X^{z_2}$, $(p \nmid x_2)$. Then as $(X) \neq (Y)$, $p \nmid x_2$. From this, by applying (2) to $\mathfrak{G}_Y$, follows $X \notin \mathfrak{G}_Y$. Hence a least subscript k_4, $(1 \leqq k_4 \leqq k)$, exists so that $(X, T_1, \cdots, T_{k_4})$ and $\mathfrak{G}_Y$ have a common non-unit element. Let Z be any such element. As $Z^p \epsilon (X, T_1, \cdots, T_{k_4-1})$, X being regarded as T_0, and $Z^p \epsilon \mathfrak{G}_Y$, follows $Z^p = E$, because, by the minimum assumption, $(X, T_1, \cdots, T_{k_4-1}) \cap \mathfrak{G}_Y = (E)$. Therefore, each of $T_1, \cdots, T_{k_4}$ being by hypothesis of order greater than p,

$$T_m \notin \mathfrak{G}_Y, \qquad 1 \leqq m \leqq k_4. \tag{4}$$

Applying the lemma first in $\mathfrak{G}_X$ and then in $\mathfrak{G}_Y$

$$Z = X^{r_0} S_{n_1}^{r_1} \cdots S_{n_{k_5}}^{r_{k_5}}, \qquad n_1 < n_2 < \cdots < n_{k_5};\ p \nmid r_{k_5},$$

$$Z = Y^{s_0} S_{m_1}^{s_1} \cdots S_{m_{k_6}}^{s_{k_6}}, \qquad m_1 < m_2 < \cdots < m_{k_6};\ p \nmid s_{k_6},$$

where $S_{n_1}, \cdots, S_{n_{k_5}}$ are members of the set $T_1, \cdots, T_{k_4}$ and $S_{m_1}, \cdots, S_{m_{k_6}} \in \mathfrak{G}_Y$. Now n_{k_5} cannot reduce to zero for then $X \in (Z) \leqq \mathfrak{G}_Y$, contrary to what we have proved; moreover by (3), $n_{k_5} > a$; from (4), one of n_{k_5}, m_{k_6}, say n_{k_5}, is greater than the other and so $n_{k_5} - 1 \geqq m_{k_6}$. Then from the two expressions for Z and the fact that X, $Y \in (S_1, \cdots, S_a)$ follows $S_{n_{k_5}} \in (S_1, \cdots, S_a, \cdots, S_{m_{k_6}}, \cdots, S_{n_{k_5}-1})$, thus contradicting the hypothesis that $S_1, \cdots, S_n$ defines a composition series for $\mathfrak{G}$. Thus our assumption that the theorem is false cannot be true and the result is proved.

THEOREM D. *If in the set $S_1, \cdots, S_n$ we replace the element S_1 by an element T of the group $\mathfrak{G}$ so that the resulting set forms a coverage system for $\mathfrak{G}$, then $(S_1) = (T)$.*

PROOF. We first prove the lemma:

If $S_1, \cdots, S_n$ be any set defining a composition series for $\mathfrak{G}$, $S_2, \cdots, S_n$ cannot be a coverage system.

This is clear for cyclic groups. We prove it by induction, assuming it true for all groups with composition series of length less than $n+1$. Let us assume it false for $\mathfrak{G}$, that is, $S_2, \cdots, S_n$ is a coverage system. Then using (1) and applying Theorem B, we see $S_1, \cdots, S_{a-1}, S_{a+1}, \cdots, S_n$ defines a composition series for $\mathfrak{G}/(X)$ and $S_2, \cdots, S_{a-1}, S_{a+1}, \cdots, S_n$ defines a coverage system. As $\mathfrak{G}/(X)$ has composition series of length n this contradicts the induction assumption and so the lemma is proved.

The proof of the theorem itself is similar. It is clear for cyclic groups. We make the induction assumption that it is true for all groups with composition series of length less than $n+1$.

The set $X, T_1, \cdots, T_k$ defines a composition series for the group $\mathfrak{G}_X$; by (3) S_a is different from each of $T_1, \cdots, T_k$. With the help of these facts and (1), Theorem B may be easily generalized to show (1) that the restclasses defined by $S_1, \cdots, S_n$ after the omission of S_a, $T_1, \cdots, T_k$ define a composition series for the factor group $\mathfrak{G}/\mathfrak{G}_X$ and (2) that the restclasses defined by $T, S_2, \cdots, S_n$ after the omission of $S_a, T_1, \cdots, T_k$ define a coverage system for $\mathfrak{G}/\mathfrak{G}_X$. None of the restclasses remaining from the set $S_2, \cdots, S_n$ has order p

in $\mathfrak{G}/\mathfrak{G}_X$, for otherwise the set $X, T_1, \cdots, T_k$ would not be a greatest possible set. Then as $\mathfrak{G}/\mathfrak{G}_X$ has composition series of length less than $n+1$, we have by the induction assumption that T and S_1 generate the same group within $\mathfrak{G}/\mathfrak{G}_X$. Therefore $T = S_1^{x_1}U$, $(p \nmid x_1,\ U \in \mathfrak{G}_X)$.

Now let $Y = S_1X$. As $S_1 \in (S_{a-1})$, we have from the definition of a $(S_{a-1}, S_a) = (S_{a-1}, X) = (S_{a-1}, Y)$. Hence Y satisfies the same conditions as X and we may therefore build a group $\mathfrak{G}_Y$ analogous to $\mathfrak{G}_X$. Exactly as before we have $T = S_1^{x_2}V$, $(p \nmid x_2,\ V \in \mathfrak{G}_Y)$. Equating the two expressions for T, $S_1^{x_2}V = S_1^{x_1}U$; whence $(XS_1)^{x_2-x_1}V = X^{x_2-x_1}U$, that is, $Y^{x_2-x_1}V = X^{x_2-x_1}U$. Now the right-hand member of this last expression is within $\mathfrak{G}_X$, while the left-hand member is within $\mathfrak{G}_Y$. But as $(X) \neq (Y)$ by Theorem C, $\mathfrak{G}_X \cap \mathfrak{G}_Y = (E)$ and consequently $UX^{x_2-x_1} = E$. Then $U = X^{x_1-x_2}$ and $T = X^{x_1-x_2}S_1^{x_1}$.

Let us assume $p \nmid x_1 - x_2$. Accordingly we may obtain an expression for X in terms of T and S_1 which we substitute into (1). By using $S_1 \in (S_r)$, we thus obtain $S_a = T^{x_3}S_r^{x_4}$ and, as the order of S_a is greater than p, $p \nmid x_4$. Then by Theorem B the restclasses containing $S_1, \cdots, S_{a-1}, S_{a+1}, \cdots, S_n$ define a composition series for $\mathfrak{G}/(T)$, while those containing $T, S_2, \cdots, S_{a-1}, S_{a+1}, \cdots, S_n$ define a coverage system. As T is within the unit restclass it is redundant, and we thus arrive at a contradiction to the lemma. Hence $x_1 - x_2$ must be divisible by p and then, as $X^p = E$, we have $(T) = (S_1)$, which proves the theorem.

THEOREM 1. *$S_1, \cdots, S_n$ defines a composition series for the group $\mathfrak{G}$ while $S_1, \cdots, S_{k-1}, T, S_{k+1}, \cdots, S_n$ is a coverage system for $\mathfrak{G}$. Then $S_1, \cdots, S_{k-1}, T, S_{k+1}, \cdots, S_n$ after a possible rearrangement of order defines a composition series for $\mathfrak{G}$.*

PROOF. For cyclic groups the result is evident. We prove the general result by induction according to the length of the composition series, assuming it true for all groups with composition series of length less than $n+1$. If the system $S_1, \cdots, S_{k-1}, S_{k+1}, \cdots, S_n$ contain no element of order p, then, as $S_1^p = E$, $k = 1$. Hence by the previous theorem $(T) = (S_1)$, from which the result follows immediately. Therefore let the system contain an element S_m where $S_m^p = E$. By Theorem B, $S_1, \cdots, S_{m-1}, S_{m+1}, \cdots, S_n$ defines a composition series for the factor group $\mathfrak{G}/(S_m)$ while $S_1, \cdots, S_{k-1}, T, S_{k+1}, \cdots, S_{m-1}, S_{m+1}, \cdots, S_n$ is a coverage system. But as $\mathfrak{G}/(S_m)$ has composition series of length n it follows from the induction assumption that the restclasses defined by this latter system after a possible rearrangement of order define a composition series for the group $\mathfrak{G}/(S_m)$, which may be written in the form

$$(\overline{S}_1, \cdots, \overline{S}_{n-1}, S_m)/(S_m), (\overline{S}_1, \cdots, \overline{S}_{n-2}, S_m)/(S_m), \cdots,$$
$$(\overline{S}_1, S_m)/(S_m), (S_m)/(S_m),$$

where the $\overline{S}_1, \cdots, \overline{S}_{n-1}$ are merely a rearrangement of the terms $S_1, \cdots, S_{k-1}, T, S_{k+1}, \cdots, S_{m-1}, S_{m+1}, \cdots, S_n$. As $S_m^p = E$, it follows from the second isomorphism theorem that

$$(\overline{S}_1, \cdots, \overline{S}_{n-1}, S_m), (\overline{S}_1, \cdots, \overline{S}_{n-2}, S_m), \cdots, (\overline{S}_1, S_m), (S_m), (E)$$

is a composition series for $\mathfrak{G}$, which establishes the result.

THEOREM E. *A coverage system of a group of rank 2 has either an element of order p or two elements S_1, S_2 for which $S_2 = S_1^z S$, $(S^p = E, p \nmid x)$.*

PROOF. For a group $\mathfrak{G}$ of rank 2 let Y be any element of order p and $\mathfrak{A}$ be the set of reciprocally cyclic subgroups of $\mathfrak{G}$ which are cyclic and contain the element Y. From the definition of a coverage system we see that it must contain, for every reciprocally cyclic subgroup $\mathfrak{H}$, an element S for which $(\mathfrak{H}, S):\mathfrak{H} = p$. From the elements of the coverage system we may therefore choose a least possible subset $\mathfrak{B}$ so that for every group $\mathfrak{H}$ of $\mathfrak{A}$ an element S of $\mathfrak{B}$ exists with $(\mathfrak{H}, S):\mathfrak{H} = p$. Let S_1 be any element of $\mathfrak{B}$ of maximum order. If $S_1^p = E$, the theorem is true; let us therefore assume $S_1^p \neq E$. Now a group $\mathfrak{H}$ exists within $\mathfrak{A}$ with $(\mathfrak{H}, S_1):\mathfrak{H} = p$, for otherwise we could omit S_1 from the set $\mathfrak{B}$, contrary to the assumption that $\mathfrak{B}$ is a least possible set. Clearly $S_1^p \in \mathfrak{H}$; then, as $\mathfrak{H}$ is cyclic and $S_1^p \neq E$, $Y \in (S_1^p) \leqq \mathfrak{H}$. Let T be an element of greatest possible order with $S_1 \in (T)$. Now $\mathfrak{G}$ having rank 2, it can be deduced from the fundamental theorem on abelian groups that (T) is reciprocally cyclic. Therefore $(T) \in \mathfrak{A}$ and so an element S_2 of $\mathfrak{B}$ exists with $(T, S_2):(T) = p$. Now as the order of S_2 cannot exceed that of S_1, $S_2^p \in (S_1^p) \leqq (T)$. Hence $S_2 = S_1^z S$, where $S^p = E$, because $S_2 \notin (T)$, $S \notin (S_1)$ and so (S, S_1^p) has rank 2. First let us assume $p \mid x$. Hence (S_2, S_1^p) would have rank 2 as it would contain (S, S_1^p). Furthermore for any group $\mathfrak{F}$ of $\mathfrak{A}$ for which $(\mathfrak{F}, S_1):\mathfrak{F} = p$, we should have $S_2^p \in (S_1^p) \leqq \mathfrak{F}$ while $S_2 \notin \mathfrak{F}$ for otherwise $(S_2, S_1^p) \leqq \mathfrak{F}$ and the cyclic group $\mathfrak{F}$ would contain the noncyclic group (S_2, S_1^p), which is impossible. Therefore $(S_2, \mathfrak{F}):\mathfrak{F} = p$, and consequently we could omit the element S_1 from the set $\mathfrak{B}$, contrary to hypothesis. Thus the assumption $p \mid x$ leads to a contradiction and so x must be prime to p. The theorem is then established.

Theorem E is not true in general for groups of rank higher than 2.*

* The following is a Gegenbeispiel: $\mathfrak{G}$ is a group of type (4 4 4) generated by A, B and C. The set of elements A, B, C, ABC^2, BCA^2, AC^2B^2, AB^2C is easily verified to be a coverage system for $\mathfrak{G}$, for which Theorem E does not hold.

THEOREM 2. *A coverage system $S_1, \cdots, S_n$ for a group $\mathfrak{G}$ of rank 2 with composition series of length $n+1$ defines a composition series for $\mathfrak{G}$ after a possible rearrangement of order.*

PROOF. For cyclic groups the result is evident. We make the assumption that it is true for all groups of rank 2 with composition series of length less than $n+1$, and then prove it for the group $\mathfrak{G}$.

By Theorem E we may select from the coverage system of $\mathfrak{G}$ either an element $S=S_m$ of order p or two elements S_m, S_k for which $S_m=S_k^{\alpha}S$, $(p\nmid x, S^p=E)$. By Theorem B, the restclasses defined by $S_1, \cdots, S_{m-1}, S_{m+1}, \cdots, S_n$ build a coverage system for the factor group $\mathfrak{G}/(S)$. But this factor group has at most rank 2 and composition series of length n. Hence, by the induction assumption, the restclasses defined by $S_1, \cdots, S_{m-1}, S_{m+1}, \cdots, S$, after a possible rearrangement of order, define a composition series which may be written in the form

$$(\overline{S}_1, \cdots, \overline{S}_{n-1}, S)/(S), (\overline{S}_1, \cdots, \overline{S}_{n-2}, S)/(S), \cdots, (\overline{S}_1, S)/(S), (S)/(S),$$

$\overline{S}_1, \cdots, \overline{S}_{n-1}$ being merely another arrangement of the terms $S_1, \cdots, S_{m-1}, S_{m+1}, \cdots, S_n$. Then, using the second isomorphism theorem, we see that

$$(\overline{S}_1, \cdots, \overline{S}_{n-1}, S), (\overline{S}_1, \cdots, \overline{S}_{n-2}, S), \cdots, (\overline{S}_1, S), (S), (E)$$

is a composition series for the group $\mathfrak{G}$. But the coverage system $S_1, \cdots, S_n$ is obtained from the set $\overline{S}_1, \cdots, \overline{S}_{n-1}, S$ by replacing the element S by the element S_m. Then we have immediately from Theorem 1 that $S_1, \cdots, S_n$, after a possible rearrangement of order, defines a composition series for the group $\mathfrak{G}$. The result is then proved.

UNIVERSITY OF NEW MEXICO

AN EXAMPLE IN THE THEORY OF POINTWISE PERIODIC HOMEOMORPHISMS*

DICK WICK HALL†

A single-valued continuous transformation $T(M)=M$ is said to be *pointwise periodic*‡ provided that for every point x in M there exists an integer N_x such that $T^{N_x}(x)=x$. The smallest such integer N_x (greater than zero) is called the *period* of x under T. For each point x of M the finite subset of M consisting of all the images of x under T is called the orbit of x under T.

In most of the familiar examples concerned with this type of transformation the limit§ of every convergent sequence of orbits consists of a set of points all having the same period. In fact the first example in which this is not the case has recently been given by G. E. Schweigert and the author.‖ In this paper we combine this example with one found by Ralph Phillips and W. L. Ayres¶ and then generalize the result to obtain what is probably the first example in which the space is a locally connected continuum while at the same time there exists a sequence of orbits converging to a set L containing points of arbitrary preassigned periods.

We let $r_1=1$ and $r_2, r_3, \cdots$ be the arbitrarily chosen periods for points in L. Then we construct a locally connected continuum M in four-dimensional euclidean (x, y, z, w) space and define a pointwise periodic homeomorphism $T(M)=M$ with the following properties:

(a) There exists a convergent sequence of orbits (G_i) under T having a limit set L which for every $i=1, 2, \cdots$ contains a free arc every interior point of which has period exactly r_i.

(b) The closure of every component of $M-L$ is a 2-cell.

---

* Presented to the Society, November 26, 1938, under the title *On pointwise periodic homeomorphisms.*

† National Research Fellow in Mathematics.

‡ See Deane Montgomery, *Pointwise periodic homeomorphisms*, American Journal of Mathematics, vol. 59 (1937), pp. 118–120.

§ The limit superior of a sequence of point sets consists of all points every neighborhood of which contains points from infinitely many sets of the sequence. The limit inferior consists of all points every neighborhood of which contains points from all but at most a finite number of sets of the sequence. If these two sets are identical, their common value is known as the limit of the sequence.

‖ See D. W. Hall and G. E. Schweigert, Duke Mathematical Journal, vol. 4 (1938), p. 723.

¶ See a forthcoming article by W. L. Ayres in Fundamenta Mathematicae. The example of Phillips and Ayres shows that a pointwise periodic homeomorphism need not be almost periodic even though the space is a locally connected continuum.

We begin by constructing the limit set L in the (x, y) plane, and do this by means of polar coordinates (r, θ). In this plane let C be the unit circle $r=1$, and define an infinite sequence of points (a_i) as follows: $a_0=(1, 0)$, $a_i=(1, [1+(i-1)/i]\pi)$, $(i\neq 0)$. Evidently, $\lim a_i=a_0$.

If the sequence (r_i), $(i=1, 2, \cdots)$, contains only a finite number of terms, we may repeat the last term infinitely many times. Hence we lose no generality in assuming this sequence to be infinite. Let A_i^0 be the arc of C from a_{i-1} to a_i in the counterclockwise direction. For each i greater than one let $A_i^1, A_i^2, A_i^3, \cdots, A_i^{r_i-1}$ be r_i-1 independent arcs (numbered inwardly) spanning* C between the points a_{i-1} and a_i. Let these arcs all lie in the sector of the circle C determined by the rays joining the origin to a_{i-1} and a_i and the portion of the circle swept out by the radius vector as it moves from a_{i-1} to a_i with increasing θ. Construct these arcs in such a way that the length of A_i^j is less than twice the length of A_i^0 for all $j\leqq r_i-1$. Also arrange that no radius of C meets any A_i^j in more than one point. We may then define A_i^n for all positive integers n by reducing the superscript n, modulo r_i. We define the limit set L by the equation

$$L=\sum_{i=1}^{\infty}\sum_{n=1}^{\infty}A_i^n.$$

For every positive integer i define

$$n_i=\sum_{j=1}^{i}2^{i-j}r_j.$$

We shall now construct an infinite sequence of point sets P_i having the following properties: (i) P_i consists of exactly n_i points, for each i; (ii) P_i lies wholly in the plane S_i: $z=1/i, w=0$; (iii) P_i converges to the set L in the plane $z=0, w=0$.

Let R_i, $(i=0, 1, 2, \cdots)$, denote the ray joining the origin to the point a_i, for each i. For i fixed consider any two rays R_{j-1} and R_j, $(j\leqq i)$. Between these rays, but not including them, construct 2^{i-j} rays equally spaced in the sector and let them be named $K_j^1, K_j^2, K_j^3, \cdots, K_j^{2^{i-j}}$ in the counterclockwise order in which they occur. The ray K_j^n will intersect L in exactly r_j points, one on each of the arcs A_j^s, $(s=0, 1, 2, \cdots, r_j-1)$. We arrange all these points of intersection, for $j=1, 2, \cdots, i$, in a linear array $w_1, w_2, \cdots, w_{n_i}$, where w_t is the point of intersection of the arc A_j^n and the ray K_j^n for the value

* An arc axb is said to span a point set M provided $(axb)M=a+b$.

$$t = n + 1 + (s - 1)r_j + \sum_{m=1}^{j-1} 2^{j-m} r_m .$$

It is easily seen that this orders the points first, according to which sector of the circle they are in, second, according to which ray they are on, and third, according to the free arc of L which contains them. Let w_t have the coordinates $(x_t, y_t, 0, 0)$. Let p_i^t be the point with coordinates $(x_t, y_t, 1/i, 0)$, and define

$$P_i = \sum_{t=1}^{n_i} p_i^t .$$

It follows easily that the sequence of point sets thus constructed satisfies (i), (ii), and (iii).

We have now reached the final stage of the construction which will be made by adding to L an infinite sequence of 2-cells, where each 2-cell is considered as the homeomorph of a triangle and its interior. It is thus meaningful to speak of the sides and vertices of a 2-cell. Every 2-cell will have one of its sides in L; this side will be called the *base* of the 2-cell. The other vertex of the 2-cell will lie in one of the planes S_i and will be referred to as the *apex* of the 2-cell. By the *altitude* we mean the maximum distance from the base to any point of the 2-cell.

For every integer i we shall construct a set of n_i 2-cells G_i^1, G_i^2, G_i^3, $\cdots$, $G_i^{n_i}$. The bases of the G_i^j will be taken as the sets

$$B_i^j = \sum_{k=1}^{i} A_k^j .$$

The apexes b_i^j will be chosen to satisfy the following conditions: (a) all the b_i^j are distinct; (b) for each (i, j), $G_i^j S_i = b_i^j$; (c) $\rho(b_i^j, p_i^j) < 1/i$. We finally require that for every $\epsilon > 0$ there exist at most a finite number of G_i^j of altitude greater than ϵ, and that the product of any of the 2-cells, which are not identical, lies in L. All of these conditions are easily satisfied in our four-dimensional space.

Our space M is now defined to be

$$M = L + \sum_{i=1}^{\infty} \sum_{j=1}^{n_i} G_i^j .$$

It is easily seen that M is a locally connected continuum.

We must now define the transformation $T(M) = M$. To do this we define $T(A_i^j) = A_i^{j+1}$ as a homeomorphism having the end points of A_i^j as fixed points. This defines T on L. For each G_i^j let $T(G_i^j) = G_i^{j+1}$

(where the superscript is reduced, modulo n_i, if necessary) be a homeomorphism agreeing with T on L and sending b_i^j into b_i^{j+1} (with the same convention on the superscripts). This defines T for every p of M. It is evident that $T(M)=M$ is a pointwise periodic homeomorphism.

If we now define

$$G_i = \sum_{j=1}^{n_i} b_i^j,$$

we see that each G_i is an orbit under T, and conditions (a) and (b) of the theorem are satisfied. The proof is thus complete.

UNIVERSITY OF PENNSYLVANIA AND
UNIVERSITY OF VIRGINIA

AN ENUMERATION OF LOGICAL FUNCTIONS

WILLIAM WERNICK

In a logical calculus of m values, abbreviated by L_m, we may deal with functions of n variables. A particular function is defined in this calculus if we assign a constant value, which may be any arbitrary one of the m possible values in L_m, as the value of that function for a particular argument. It is the purpose of this note to enumerate, among all functions of n variables in L_m: those which depend on all n variables in the argument; those which depend on just $(n-1)$ of the variables in the argument, being independent of one of them; and so on; finally those which are completely independent of all the variables in the argument.

Since each variable in the argument may assume values from $1, \cdots, m$, independently, there are m^n possible arguments, and since to each argument we may assign independently, as a functional value, any of the m values $1, \cdots, m$, there are in all m^{m^n} possible functions of n variables.

Let V_n be the total number of all functions of n variables in L_m. Then we have from the above

$$V_n = m^{m^n}. \tag{1}$$

Let U_{nk} be the number of functions of n variables which depend on exactly k of them. (It is this expression for which we are seeking an explicit evaluation.) Since k variables may be selected from n of them in just $C_{n,k}$ ways, we have the relation:

$$U_{nk} = C_{n,k}U_{kk}. \tag{2}$$

Since V_n includes all functions of n variables in L_m, we may write V_n as the sum of: all functions of n variables which depend on all of them; all functions of n variables which depend on all but one of them; and so on; that is, $V_n = \sum_{k=0}^{n} U_{nk}$, which, by the relation above, becomes

$$V_n = \sum_{k=0}^{n} C_{n,k} U_{kk}. \tag{3}$$

This leads to the following implicit recursive definition for U_{kk}:

$$V_0 = U_{00}, \tag{4_0}$$

$$V_1 = C_{1,0} U_{00} + C_{1,1} U_{11}, \tag{4_1}$$

$$V_2 = C_{2,0} U_{00} + C_{2,1} U_{11} + C_{2,2} U_{22}, \tag{4_2}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots,$$

$$V_k = C_{k,0} U_{00} + C_{k,1} U_{11} + \cdots + C_{k,k} U_{kk}. \tag{4_k}$$

We may solve this system of $k+1$ equations in the $k+1$ unknowns $U_{00}, U_{11}, \cdots, U_{kk}$ by the use of Cramer's rule and some manipulation of determinants to obtain the value for U_{kk}:*

$$U_{kk} = \sum_{i=0}^{k} (-1)^i C_{k,i} V_{k-i}, \tag{4}$$

which by application of (1) gives the explicit value we want for U_{kk}:

$$U_{kk} = \sum_{i=0}^{k} (-1)^i C_{k,i} m^{m^{k-i}}. \tag{5}$$

Finally, using this result in formula (2), we have the actual enumeration formula; that is, the number of functions of n variables in L_m which depend on exactly k of them is given by the formula:

$$U_{nk} = C_{n,k} \sum_{i=0}^{k} (-1)^i C_{k,i} m^{m^{k-i}}. \tag{6}$$

REMARK. We may, in formula (4), make use of the apparent analogy with the binomial expansion. If we agree to interpret, that is, symbolize, V_i as v^i, where the i in v^i is an exponent, and write this interpretation $V_i \equiv v^i$, we have, for example, $V_0 \equiv v_0 = 1$.† Then our

* I am indebted for the actual carrying out of this solution to Dr. J. C. C. McKinsey. The solution, though elementary, is quite lengthy and may be left as an exercise for the reader.

† We write $\equiv$ to mean "symbolically equals," and $=$ to mean "equals." These are not interchangeable, since V_0 is not equal to 1, but to m.

equation (4) becomes

$$U_{kk} \equiv \sum_{i=0}^{k} C_{k,i}(-1)^i v^{k-i}. \tag{4'}$$

But this expansion on the right is simply the binomial expansion for $(v-1)^k$, so our simplified formula now is

$$U_{kk} \equiv (v-1)^k. \tag{7}$$

For the term U_{00} we get $U_{00} \equiv (v-1)^0 = 1 = v^0 \equiv V_0$ as required in (4_0). We get corresponding definitions for the others:

$$U_{11} = V_1 - V_0, \tag{7_1}$$

$$U_{22} = V_2 - 2V_1 + V_0, \tag{7_2}$$

. . . . . .

Each of these is made explicit by the use of formula (1). The simplified formula for U_{nk} is now

$$U_{nk} \equiv C_{n,k}(v-1)^k. \tag{8}$$

To evaluate a particular term, expand the right side, and apply the definitions: $v^0 \equiv V_0 = m^{m^0}$, $v^i \equiv V_i = m^{m^i}$.

APPLICATION. If L_m is the two valued calculus of sentences, and if we consider truth functions of two variables $f_j(p, q)$, we have $m=2$ and $n=2$. Then $V_2 = 2^{2^2} = 16$. There are 16 possible truth functions of two variables in L_m. Of these 16 functions, there are $U_{20} = C_{2,0}2^{2^0} = 2$ functions that depend on 0 variables, that is, that are independent of both p and q ("$f(p, q)$ is always true," "$g(p, q)$ is always false"). There are $U_{21} = C_{2,1}(2^{2^1} - 2^{2^0}) = 4$ functions that depend on one variable only, that is, on p alone or on q alone, but not on both ("p is true," "p is false," "q is true," and "q is false"). There are $U_{22} = C_{2,2}(2^{2^2} - 2 \cdot 2^{2^1} + 2^{2^0}) = 10$ functions that depend on both variables (all the remaining functions in V_2, including conjunction, disjunction, implication, and so on).

NEW YORK UNIVERSITY

THE UNIVERSALITY OF FORMAL POWER SERIES FIELDS*

SAUNDERS MacLANE

In a recent paper,† André Gleyzal has constructed ordered fields consisting of certain "transfinite real numbers" and has established the interesting result that any ordered field can be considered as a subfield of one of these transfinite fields. These fields prove to be identical with fields of formal power series in which the exponents are allowed to range over a suitable ordered abelian group. Such fields were first introduced by Hahn,‡ while they have been analyzed in terms of generalized valuations by Krull.§

Gleyzal applied his construction of transfinite real numbers not only to the case when the coefficient field consisted of real numbers, but also to suitable fields of characteristic p. He conjectured that this construction should yield a "universal" field of characteristic p. We show here that Krull's technique can be used to establish Gleyzal's conjecture.

1. **Formal power series.** If K is any field and Γ any ordered abelian group (its order may be non-archimedean), we form all power series $x=\sum a_\alpha t^\alpha$ with coefficients a_α in K, exponents α in Γ, and summed over a subset N of elements α from Γ which is normally ordered by the given linear order in Γ. Such a series could also be written as

$$x = a_{\alpha_1}t^{\alpha_1} + a_{\alpha_2}t^{\alpha_2} + \cdots + a_{\alpha_\rho}t^{\alpha_\rho} + \cdots , \tag{1}$$

summed over all ordinal numbers ρ less than a fixed σ, and with exponents $\alpha_1<\alpha_2<\cdots<\alpha_\rho<\cdots$ increasing monotonically. The product of two formal powers t^α and t^β is defined as $t^{\alpha+\beta}$, where $\alpha+\beta$ is the sum in the group Γ. On this basis, the usual formal definitions of multiplication and addition make the set of all series (1) a field, which we denote by $K\{t^\Gamma\}$.

Hahn also gave a similar construction for a non-archimedean ordered group from a given ordered group G (say an additive group of

* Presented to the Society, April 8, 1939.

† A. Gleyzal, *Transfinite numbers*, Proceedings of the National Academy of Sciences, vol. 23 (1937), pp. 581–587.

‡ H. Hahn, *Über die nichtarchimedischen Grössensysteme*, Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften, Vienna, section IIa, vol. 116 (1907), pp. 601–653.

§ W. Krull, *Allgemeine Bewertungstheorie*, Journal für die reine und angewandte Mathematik, vol. 167 (1931), pp. 160–196.

real numbers) and any ordered set B. One simply chooses for each b in B a "basis" element e_b, and then forms all sums $\alpha=\sum g_b e_b$, with g_b in G, taken over the indices b of a subset M of B which is well ordered by the given order of B. The formal sum of two such expressions α is again a similar expression. Under the usual lexicographic ordering, all these expressions constitute an ordered group $G\{e_B\}$.

In particular, one may start with the ordered (additive) group R of all real numbers, construct the group $\Gamma=R\{e_B\}$ and thence the field $K=R'\{t^\Gamma\}$, where R' is the real number field. This particular power series field is isomorphic to Gleyzal's field of "transfinite real numbers* with base order B."

Hahn also showed how the group Γ could be redistilled from the ordered field $K=R'\{t^\Gamma\}$ by considering each element in Γ as a set of all elements in K of the same "order of magnitude." If this process is applied again, the orders of magnitude in Γ yield the original† ordered set B.

2. **Algebraic closure.** A *root group* Γ is one with the property that for every integer n and every α in Γ there is a γ in Γ with $n\gamma=\alpha$. The particular group $R\{e_B\}$ is a root group.

THEOREM 1. *The power series field $K\{t^\Gamma\}$ is algebraically closed if the coefficient field K is algebraically closed and if the ordered abelian group Γ of exponents is a root group.*

PROOF. In the power series field $S=K\{t^\Gamma\}$ we introduce a valuation V by setting $V(x)=\alpha_1$ if $a_{\alpha_1}t^{\alpha_1}$ is the first nonvanishing term in the power series (1) for x. In this valuation, Γ is the value group and K the field of residue classes. Furthermore,‡ S is maximal with respect to this valuation, in the sense that any proper extension $S'>S$ to which the valuation V has been extended must either have a larger value group or a larger residue class field than S.

Suppose now that S is not algebraically closed, so that S has a proper finite normal extension N. Certainly S is (algebraically) perfect, so that N/S is separable. The valuation V of S can be extended to N by the usual methods, for S is (topologically) perfect§ with respect to V. The ordinary Newton polygon construction shows that

* To establish the isomorphism, observe that Gleyzal's product forms Π_μ form a multiplicative group isomorphic to $R\{e_B\}$, and that his transfinite real numbers or sum forms are manipulated exactly as are the formal series (1).

† This twofold order of magnitude process yields essentially Gleyzal's base order for K.

‡ Krull, loc. cit., Theorem 26.

§ Krull, loc. cit., §9 and Theorem 27.

each element c of N has a value of the form α/n, with α in Γ. Since Γ is a root group, $\alpha/n \in \Gamma$, and Γ is thus the value group of N. On the other hand, the residue class field of N must be an algebraic extension of the algebraically closed residue class field K of S. Thus N presents a proper extension of S in which neither value group nor residue class field is extended, contrary to the maximal property of S.

3. **Universal fields.** We call a field F *universal* if every other field F' which has the same cardinal number and the same characteristic as F is isomorphic to a subfield of F.

THEOREM 2. *A nondenumerable field F is universal if and only if it contains an algebraically closed subfield F_0 which has the same cardinal number as F.*

PROOF. We need only provide a map of any F' into F. As in the Steinitz theory, let T_0, T' be respectively transcendence bases for F_0, F' over the prime subfield P. By the axiom of choice, the cardinal number of T_0 is some aleph. Since P is at most denumerable, the usual computations with cardinal numbers then show that $P(T_0)$ and its algebraic extension F_0 must have the same cardinal number as does T_0. Since F_0 and F' have the same cardinal number, we conclude that T_0 and T' have the same cardinal number, so we can set T_0 and T' into one-to-one correspondence. This correspondence gives an isomorphism of $P(T')$ to $P(T_0)$. Because F is algebraically closed, this isomorphism can be extended* from $P(T')$ to its algebraic extension F'. We have proved F_0 and hence F universal.

THEOREM 3. *If the ordered abelian root group Γ contains an element different from 0, while the coefficient field K is algebraically closed, then the power series field $K\{t^\Gamma\}$ is universal.*

This field $K\{t^\Gamma\}$ includes all ordinary power series in t^α, with $\alpha \neq 0$ in Γ, hence has at least the power of the continuum. Theorem 3 thus follows from Theorems 1 and 2. In particular, this theorem yields relatively simple universal fields whose cardinal numbers are that of the continuum; for example, the fields $K\{t^\Gamma\}$ with K the absolutely algebraic algebraically closed field of any characteristic and Γ the group of all rational numbers. This theorem also includes Gleyzal's conjectured fields, where K is absolutely algebraic and Γ is any one of the groups $R\{e_B\}$ of §1.

HARVARD UNIVERSITY

* B. L. van der Waerden, *Moderne Algebra*, §60.

THEOREMS ASSOCIATED WITH THE RIESZ AND THE DIRICHLET'S SERIES METHODS OF SUMMATION

H. L. GARABEDIAN

1. **Introduction.** It is the object of this paper to establish several theorems suggested by a certain theorem due to Hardy.* Hardy's theorem need not be stated here since it is a special case of Theorem 1 of this paper.

We are concerned with the infinite series

$$\sum_{n=1}^{\infty} u_n = u_1 + u_2 + \cdots, \tag{1.1}$$

where we define $s_n = u_1 + u_2 + \cdots + u_n$.

We list the several different methods of summation with which this account is occupied.

A. *The Riesz definition.*† We write $C_\lambda^r(\omega) = \sum_{\lambda_n < \omega} (\omega - \lambda_n)^r u_n$, $(r > 0)$, where $\{\lambda_n\}$ is a sequence of real increasing numbers whose limit is infinite and such that $\lambda_1 \geqq 0$. If $\omega^{-r} C_\lambda^r(\omega) \to U$ as $\omega \to \infty$, the series (1.1) is said to be summable (λ, r) to the sum U. If in the general definition we put $r = 1$, $\omega = \lambda_n$, we obtain

$$(\mu_1 s_1 + \mu_2 s_2 + \cdots + \mu_{n-1} s_{n-1})/\lambda_n \tag{1.2}$$

where $\mu_i = \lambda_{i+1} - \lambda_i$ and $\lambda_n = \mu_1 + \mu_2 + \cdots + \mu_{n-1}$, $\lambda_1 = 0$. We refer to means of the type (1.2) as Riesz *means*, as distinguished from the Riesz *typical means* of the general definition, and designate them henceforth by the symbol $(\lambda, 1)$. This is the natural generalization of Cesàro's first mean which suggests itself in the attachment of varying weights to the successive partial sums s_ν.

B. *The Dirichlet's series definitions.*‡ A series (1.1) is said to be summable by the Dirichlet's series method provided that

$$\lim_{s \to 0} \sum_{n=1}^{\infty} u_n e^{-\nu_n s}$$

exists, where $\{\nu_n\}$ is a sequence of positive increasing real numbers whose limit is infinite, and where the Dirichlet's series converges when $\Re(s) > 0$, s being restricted to this half plane.

* G. H. Hardy, Proceedings of the London Mathematical Society, (2), vol. 8 (1910), pp. 301–320, p. 311.

† M. Riesz, Comptes Rendus, vol. 149 (1909), pp. 909–912.

‡ See, for example, H. L. Garabedian, Annals of Mathematics, (2), vol. 32 (1931), pp. 83–106, p. 85.

2. **Extensions of Hardy's theorem.** We propose to establish the following theorems.

THEOREM 1. *Suppose that* $\lim_{n\to\infty} s_n e^{-\nu_n s} = 0$ *for all values of* s *for which* $\Re(s) > 0$, *that the series* (1.1) *is summable* $(\lambda, 1)$ *to the sum* U, *and that* ν_n *is a logarithmico-exponential* function of* λ_n *such that* $\nu_n = O(\lambda_n^{\Delta})$, *where* Δ *is any positive constant. Then, the series* $\sum u_n e^{-\nu_n s}$ *converges for* $\Re(s) > 0$ *and*

$$\lim_{s\to 0} \sum_{n=1}^{\infty} u_n e^{-\nu_n s} = U.$$

THEOREM 2. *The Dirichlet's series definitions of summability include* $(\lambda, 1)$ *summability provided that* ν_n *is a logarithmico-exponential function of* λ_n *which tends to infinity with* n *but not as slowly as* log n *nor faster than* λ_n^{Δ}, *where* Δ *is any positive constant however large.*

We shall understand in both theorems that $s = \sigma + i\tau$ approaches zero over a point set lying within an angle with vertex at the origin such that

$$|\operatorname{am} s| \leqq a < \pi/2.$$

Theorem 1 reduces to Hardy's theorem for the case that $\nu_n = \lambda_n$. We observe further that the restriction that $\nu_{n+1} - \nu_n \to 0$ as $n \to \infty$ of Hardy's theorem has been eliminated in Theorem 1.

Theorem 2 eliminates the initial restriction of Theorem 1 and thus affords a regulation inclusion theorem. Theorem 2 recalls the following theorem previously established by the author.†

THEOREM 3. *The Dirichlet's series definitions of summability include* (C, r) *summability provided that* ν_n *is a logarithmico-exponential function of* n *which tends to infinity with* n *but not as slowly as* log n *nor faster than* n^{Δ}, *where* Δ *is any positive constant however large.*

This theorem is at once slightly more general and slightly less general than Theorem 2. To understand this apparent contradiction we need to make use of the following theorem due to Hardy.‡

THEOREM 4. *If the series* (1.1) *is summable* (λ, r) *to the sum* U, *and if* ν_n *is a logarithmico-exponential function of* λ_n *such that* $\nu_n = O(\lambda_n^{\Delta})$, *where* Δ *is any positive constant, then the series* (1.1) *is summable* (ν, r) *to the sum* U.

* Hardy, Proceedings of the London Mathematical Society, (2), vol. 15 (1916), pp. 72–88, p. 75.

† Loc. cit., footnote 3, p. 105.

‡ Loc. cit., footnote 4, p. 72.

Some examples of the equivalence and inclusion* relations implied by this theorem are

$$(\log_3 n, 1) \supset (\log_2 n, 1) \supset (\log n, 1) \approx ([\log n]^2, 1) \supset (n, 1)$$

$$(2.1) \qquad = (C, 1) \approx (n^r, 1) \approx (n \log n, 1) \supset (k^n, 1), \qquad r > 0,\ k > 1.$$

Incidentally, summability $(k^n, 1)$, for $k>1$, is equivalent to convergence.†

It is also pertinent to this discussion that none of the $(\lambda, 1)$ methods can include summability (C, r) for $r>1$.‡ On the other hand, it follows from an example given by Hardy§ that summability (C, r), for r arbitrarily large, cannot include summability $(\log n, 1)$.

It follows as a result of Theorem 3 that all of the Dirichlet's series methods of summation which include (C, r) summability include all of the $(\lambda, 1)$ methods equivalent to $(C, 1)$ summability. Theorem 2 involves a wider class of $(\lambda, 1)$ methods than Theorem 3 but implies nothing as to (C, r) summability for $r>1$.

The particular method of Riesz means, summability $(\log n, 1)$, has been designated as *logarithmic summability*,‖ and has been the source of a number of investigations relating to the summability and convergence of slowly oscillating series. Thus, it is of especial interest that Theorem 2 affords a means of exhibiting a class of Dirichlet's series methods of summation which include summability $(\log n, 1)$. This result is expressed in the following corollary to Theorem 2.

COROLLARY. *The Dirichlet's series methods of summation include summability* $(\log n, 1)$ *provided that* ν_n *is a logarithmico-exponential function of n which tends to infinity with n but not as slowly as* $\log n$ *nor faster than* $(\log n)^{\Delta}$, *where* Δ *is any positive constant however large.*

3. **Proofs of Theorems 1 and 2.** The proofs of Theorems 1 and 2, which for the present are carried simultaneously, are based on several theorems due to Hardy and Riesz to which we shall refer as the occasion arises.

* We understand $A \subset B$ to mean that every series summable by a method A is also summable by a method B to the same limit, or that the method B includes the method A. Further, we interpret $A \approx B$ to mean that the two methods are equivalent, that is, each method includes the other.

† Hardy and Riesz, *The General Theory of Dirichlet's Series*, Cambridge Tracts in Mathematics and Mathematical Physics, no. 18, p. 35.

‡ H. L. Garabedian and W. C. Randels, Duke Mathematical Journal, vol. 4 (1938), pp. 529–533, p. 531.

§ Loc. cit., footnote 1, p. 310.

‖ Hardy, loc. cit., footnote 4.

Since by assumption

$$\sum_{n=1}^{\infty} u_n e^{-\nu_n s} \tag{3.1}$$

is summable $(\lambda, 1)$, for $s=0$, to the sum U, it is also summable $(\nu, 1)$, for $s=0$, to the same value. This statement follows from Theorem 4.

Since (1.1) is summable $(\nu, 1)$, (3.1) is uniformly summable $(\nu, 1)$ throughout the angle $\boldsymbol{a}$ to the sum $f(s)$. This statement is due to Hardy and Riesz.*

Inasmuch as (3.1) is summable $(\nu, 1)$ to the sum U, for $s=0$, and since (3.1) is uniformly summable $(\nu, 1)$ to the sum $f(s)$ throughout the angle $\boldsymbol{a}$, $f(s)\to U$ as $s\to 0$ along any path lying entirely within the angle $\boldsymbol{a}$. This follows from another theorem due to Hardy and Riesz.†

If the hypotheses of Theorem 1 obtain and if $\{\nu_{n+1}-\nu_n\}$ is a null sequence, then the series (3.1) converges for $\Re(s)>0$.‡ The extra restriction will be eliminated by an argument given below. Now, the series (3.1) converges to the value $f(s)$ to which it is also summable $(\nu, 1)$. It follows that the value approached as $s\to 0$ must be the same in both cases. Thus, with the restrictions on ν_n which have been stated,

$$\lim_{s\to 0} \sum_{n=1}^{\infty} u_n e^{-\nu_n s} = U.$$

This establishes Theorem 1.

The abscissa of convergence of a Dirichlet's series is given by the formula§

$$\sigma_0 = \limsup_{n\to\infty} (\log |s_n|)/\nu_n.$$

Now, it is known that if the series (1.1) is summable $(\nu, 1)$ to the value U, then‖

$$s_n = o(\nu_{n+1}/(\nu_{n+1}-\nu_n)) + U.$$

Thus, it is easily verified that if the series (1.1) is summable $(\nu, 1)$, then $s_n = o(n^2)$. Accordingly, we have in this case

* Hardy and Riesz, loc. cit., footnote 8, Theorem 23.

† Hardy and Riesz, loc. cit., footnote 8, Theorem 28.

‡ Hardy, loc. cit., footnote 1, p. 311. It is assumed in Hardy's theorem that $\lambda_{n+1}/\lambda_n\to 1$. However, we note that, of the $(\lambda, 1)$ methods listed in (2.1), this restriction eliminates only the trivial case for which $\lambda_n = k^n$, $(k>1)$.

§ Hardy and Riesz, loc. cit., footnote 8, Theorem 7.

‖ Hardy and Riesz, loc. cit., footnote 8, Theorem 21.

$$\sigma_0 \leqq \limsup_{n\to\infty} (2 \log n)/\nu_n.$$

In order that $\sigma_0 \leqq 0$, in which case (3.1) will converge in the right half of the s-plane, it is sufficient that ν_n tend to infinity faster than $\log n$. The argument used to complete the proof of Theorem 2 is the same as the one used above in connection with Theorem 1.

Notice that if $\{\nu_{n+1}-\nu_n\}$ is not a null sequence, then ν_n tends to infinity faster than $\log n$. This eliminates the extra restriction used in the proof of Theorem 1.

NORTHWESTERN UNIVERSITY

ON CERTAIN IDEALS OF DIFFERENTIAL POLYNOMIALS*

J. F. RITT AND E. R. KOLCHIN

Introduction. Let Σ be an ideal of differential polynomials in the unknowns $y_1, \cdots, y_n$. If the manifold of Σ is composed of s manifolds $\mathfrak{M}_1, \cdots, \mathfrak{M}_s$ not necessarily irreducible, none of which has a solution in common with any other, Σ has a unique representation as the product of s ideals $\Sigma_1, \cdots, \Sigma_s$ whose manifolds are, respectively, the $\mathfrak{M}_i$.†

Most of the present note is concerned with decompositions of the foregoing type and considers the case in which one of the $\mathfrak{M}_i$, say $\mathfrak{M}_1$, is composed of a single solution, that is, of a set of functions $\bar{y}_1, \cdots, \bar{y}_n$ contained in the underlying field. We shall examine, for this special case, the structure of the ideal Σ_1. Details will be given only for the case of a single unknown; the extensions to several unknowns are too obvious to require explicit mention. It will suffice, furthermore, to treat the case in which $\mathfrak{M}_1$ is composed of the solution $y=0$.

In §9, we consider a problem closely related to the theorem of decomposition stated above.

1. **On the structure of Σ_1.** Let Σ be an ideal of forms in the unknown y. Let $y=0$ be an essential irreducible manifold for Σ. Let Σ be the product of Σ_1 and Σ_2 where Σ_1 has $y=0$ as its manifold and Σ_2 does not admit $y=0$ as a solution. Let p be a positive integer such that y^p is contained in Σ_1.

* Presented to the Society, September 8, 1939.

† Proceedings of the National Academy of Sciences, vol. 25 (1939), p. 90. *Product* is defined in the expected way. That the intersection of the Σ_i is identical with their product follows immediately from the fact that the Σ_i, considered as *algebraic* ideals, are *paarweise teilerfremd*. See van der Waerden, *Moderne Algebra*, vol. 2, p. 46.

We shall prove that Σ_1 *is the ideal generated by* Σ *and* y^p, that is, the intersection of all ideals containing Σ and y^p.

PROOF. Obviously Σ_1 contains the ideal generated as above. What has to be proved is the converse of this fact. Let G be any form in Σ_1. Σ_2 contains a form $1-H$, where H vanishes for $y=0$. Then Σ contains $G(1-H)$ and, therefore, $G(1-H^q)$ where q is any positive integer. Now, if q is large,

$$H^q \equiv 0, (y^p). \tag{1}$$

Let q be fixed at a value large enough for (1) to hold, and let $M=G(1-H^q)$. Then $G\equiv M$, (y^p), and this establishes the theorem.

2. **Condition for p to be unity.** We are going to examine now the case in which Σ contains a form of the type $y+A$, where A, considered as a polynomial in the y_j, has no term of degree less than 2. Ideals of this type form a natural and interesting class; a very special example is the ideal generated by y_1^2-4y. We are going to prove that the p of §1 may be taken as unity. That is, Σ_1 *consists of all forms which vanish for* $y=0$.

What we have to prove, of course, is that Σ_1 contains y. Our procedure will be as follows. For some p, y^p is in Σ_1. Then, for any i, $y_i^{2^i p}$ is in Σ_1. Let $F=y+A$ be the form mentioned in the statement of our theorem. Then

$$y \equiv -A, (F). \tag{2}$$

We shall subject (2) to an iterative process and derive a relation $y\equiv K$, (F), where every term of K contains some $y_i^{2^i p}$ as a factor, i depending on the term. This will establish the theorem.

3. **Bound on degrees.** Let a form P in y be of degree g in some y_j, $j\geqq 0$. We shall show that P', the derivative of P, has a degree in y_j which does not exceed $g+1$. For let L, any term of P, be divisible by y_j^q with $q\leqq g$, and by no higher power of y. Let $L=y_j^q M$. We have, indicating first derivatives by an accent,

$$L' = qy_j^{q-1}y_{j+1}M + y_j^q M'.$$

M' consists of a set of terms, one of which will be divisible by the first power of y_j if M involves y_{j-1}. This is enough to prove our statement.

4. **The first substitution.** Let us suppose that, in addition to (2), we have a second relation $y\equiv B$, (F). In the second member of (2), let y be replaced by B and each y_j by the jth derivative of B. Then $-A$ goes over into a form C. It is easy to see that $y\equiv C$, (F).

Let r be a positive integer which is not less than the order of A in y. Let g be an integer, exceeding unity, such that each term of A is of total degree not less than g in $y, \cdots, y_r$. In A, we replace y_j by $-A^{(j)}$, $j=0, \cdots, r$, superscripts indicating differentiation.* Then $-A$ goes over into a form A_1 and $y \equiv A_1$, (F). Each term in A_1 is of total degree not less than g^2 in $y, \cdots, y_{2r}$. By §3, the $A^{(j)}$, $j \leqq r$, are of degree not greater than r in any one of the letters $y_{r+1}, \cdots, y_{2r}$.

Let L be a term in A_1, of total degree $d \geqq g^2$ in the y_j. The power product of degree d in L is the product of a set of power products taken from the $A^{(j)}$. If M is any of the latter power products, the total degree of M is at least g, hence at least g/r times the degree of M in any one of $y_{r+1}, \cdots, y_{2r}$. Thus, the degree of L in any one of $y_{r+1}, \cdots, y_{2r}$ is not more than $(rd)/g$.

5. **The second substitution.** Differentiating A_1, we consider the $A_1^{(j)}$ for $j=0, \cdots, r$. No $A_1^{(j)}$ is of degree exceeding r in any y_i with $2r < i \leqq 3r$. Let L, of some total degree d, be a term in some $A_1^{(j)}$. Then L, since it is derived from a term of total degree d in A_1, is of degree not exceeding $rdg^{-1}+r$ in any y_i with $r < i \leqq 2r$. As $d \geqq g^2$, we have $rdg^{-1}+r \leqq rd(g^{-1}+g^{-2})$.

In the second member of (2), we replace each y_j by $A_1^{(j)}$. We find a relation $y \equiv A_2$, (F), with each term of A_2 of total degree at least g^2. If L is a term in A_2, of some total degree d, the degree of L in any y_i with $2r < i \leqq 3r$ does not exceed rdg^{-2} and the degree of L in any y_i with $r < i \leqq 2r$ does not exceed $rd(g^{-1}+g^{-2})$.

6. **Continuation.** In the third step, we substitute the $A_2^{(j)}$ into (2). After t steps, we have $y \equiv A_t$, (F), with each term in A_t of total degree at least g^{t+1}. Let L be a term in A_t of some total degree d. Let j be any positive integer not greater than t. Then the degree of L in any y_i with $jr < i \leqq (j+1)r$ does not exceed

$$rd(g^{-j} + g^{-j-1} + \cdots + g^{-t}) < 2rdg^{-j}.$$

7. **Completion of proof.** Let t of §6 be the square of a positive integer s. The total degree of L of §6 in the y_i with $i > sr$ is no more than

$$2rd(rg^{-s} + \cdots + rg^{-t}) < 4r^2dg^{-s}.$$

Let s be so great that

$$4r^2g^{-s} < 1/2.$$

Then the total degree of L in the y_i with $i \leqq sr$ is at least $d/2$. Thus, for some particular y_i with $i \leqq sr$, the degree of L in y_i is at least

* $A^{(0)} = A$.

$$\frac{d}{2(rs+1)} \geqq \frac{g^s+1}{2(rs+1)}. \tag{3}$$

We refer now to §2. If s is large, the second member of (3), if $i \leqq sr$, will exceed $2^i p$. This completes the proof of our theorem.

8. **Higher values of p.** It is not an unnatural conjecture that, if Σ contains a form $y^n + A$ with every term in A of degree greater than n, p of §1 may be taken as n. We give an example to show that the least p may exceed n.

Let Σ be the ideal generated by $F = y^3 + y_1^4$. If Σ_1 contained y^3, there would exist a relation

$$y^3(1 - H) = MF + M_1F' + \cdots + M_rF^{(r)}, \tag{4}$$

with H vanishing for $y = 0$. For the second member of (4) to yield the term y^3 which the first member contains, it would be necessary for M to have unity as one of its terms. Then MF would have y_1^4 as a term. Equating terms of degree 4 and weight 4 for both sides of (4), we would find $y_1^4 \equiv 0, (y^3)$, which is readily shown to be false.

9. **A generalization.** Let F and A be two forms in $y_1, \cdots, y_n$, both of class n and algebraically irreducible. Suppose that the general solution $\mathfrak{M}$ of A is contained in the manifold of F and is essential in that manifold. It is known how the essentiality of $\mathfrak{M}$ is reflected in the structure of F.* Suppose now that S^tF, where S^t is as in the indicated theorem of structure, has a term C_jA. It can be shown, by the method of §§2–7 above, that *there exists a relation* $AH \equiv 0, (F)$, *where* H *does not hold* $\mathfrak{M}$.

COLUMBIA UNIVERSITY

* American Journal of Mathematics, vol. 60 (1938), p. 14.

A THEOREM ON SIMULTANEOUS REPRESENTATION OF PRIMES AND ITS COROLLARIES*

ARNOLD E. ROSS

1. **Simultaneous representation of primes.** Two numbers m and M are said to be represented simultaneously by a ternary form

$$f = ax^2 + by^2 + cz^2 + 2ryz + 2sxz + 2txy \tag{1}$$

and its reciprocal†

$$F = AX^2 + BY^2 + CZ^2 + 2RYZ + 2SXZ + 2TXY \tag{2}$$

if there exist integers x, y, z and X, Y, Z such that $f(x, y, z) = m$, $F(X, Y, Z) = M$ and $xX + yY + zZ = 0$.

The case of interest is that in which representation is not only simultaneous but also proper.‡ One is usually interested in the existence of such numbers m and M, fulfilling certain conditions, with the view of a suitable normalization of the given form f and its reciprocal F.§

In this paper we will require that m and M be a pair of simultaneously and properly represented distinct odd primes or doubles of such primes and derive a normalized form permitting some interesting applications. We note that the first coefficient a of f and the third coefficient C of F are represented simultaneously and properly and express our result as the following theorem.

THEOREM 1. *If f is a ternary quadratic form with a properly primitive reciprocal and if f is* (i) *properly or* (ii) *improperly primitive, then it is equivalent to a form f' such that* (i) *a' and C' are distinct odd primes not dividing $2\Omega\Delta$, or* (ii) *$a' = 2\alpha$ and α and C' are distinct odd primes not dividing $2\Omega\Delta$. Here a' is the leading coefficient of f', and C' is the third coefficient of the reciprocal F' of f'.*

We note that since F is properly primitive it represents properly an integer prime to any assigned integer and hence to $2\Omega\Delta$. If $\Omega\Delta$ is odd, then F represents properly an integer congruent to 1 (mod 4)

* Presented to the Society in part, April 9, 1937, under the title *On certain rational transformations.*

† See Dickson, *Studies in the Theory of Numbers*, University of Chicago Press, p. 12.

‡ Ibid.

§ Dickson, ibid., pp. 15–17 and 54–60; P. Bachman, *Die Arithmetik der quadratischen Formen*, vol. 1, p. 64; H. J. S. Smith, *Collected Works*, vol. 1, pp. 455–509.

and one congruent to 3 (mod 4). Hence we may assume that C is prime to $2\Omega\Delta$ and, if $\Omega\Delta$ is odd,

$$C \equiv \Omega \pmod 4. \tag{3}$$

Then the binary form

$$\psi = ax^2 + 2txy + by^2, \qquad ab - t^2 = \Omega C, \tag{4}$$

is properly or improperly primitive according as f is properly or improperly primitive.

For, in view of a well known property of determinants,

$$aS + tR + sC = 0, \quad tS + bR + rC = 0, \quad sS + rR + cC = \Omega\Delta. \tag{5}$$

The g.c.d. g of a, t, b divides $ab-t^2=\Omega C$ and also the determinant $D=\Omega^2\Delta$ of f. Since C is prime to $2\Omega\Delta$, g divides Ω and is prime to C. Then by (5_1) and (5_2), g divides sC and rC, and hence s and r. Therefore, since g divides Ω, it divides cC and hence c by (5_3). Thus g divides the g.c.d. of the coefficients $a, \cdots, t$ of f, and, since f is primitive, $g=1$. Hence ψ is primitive.

If f is properly primitive, so is ψ. The determinant of f is $abc+2rst-ar^2-bs^2-ct^2=\Omega^2\Delta$. Let $\Omega^2\Delta$ be even. If ψ were improperly primitive, then c would be even by the above and f would be improperly primitive. If $\Omega\Delta$ is odd, $\Omega C\equiv 1 \pmod 4$ by (3), whereas if ψ were improperly primitive, we would have $\Omega C=ab-t^2\equiv -1 \pmod 4$.

If f is improperly primitive, ψ is also improperly primitive since it is primitive.

If ψ is properly primitive, it represents properly an odd prime a_1 not dividing $2\Omega\Delta C$. If ψ is improperly primitive, it represents properly an integer $a_1=2\alpha$ such that α is an odd prime not dividing $2\Omega\Delta C$.* Hence ψ is equivalent to a form ψ_1 with a_1 as the leading coefficient. A transformation carrying ψ into ψ_1 carries f into $f_1=a_1x^2+\cdots$. The contragredient (defined by the transpose of the reciprocal of its matrix) transformation which carries F into the reciprocal F_1 of f_1, leaves C unaltered.†

The binary section $\Psi=B_1Y^2+2R_1YZ+C_1Z^2$, $(B_1C_1-R_1^2=\Delta a_1)$, of F_1 is properly primitive, since $C_1=C$ is odd and since the g.c.d. g of B_1, R_1, C_1 divides Δa_1 and hence is equal to unity by virtue of the choice of C $(=C_1)$ and a_1. Thus, as above, the properly primitive binary Ψ represents properly on odd prime C' not dividing $\Omega\Delta a_1$.‡

* H. Weber, Mathematische Annalen, vol. 20 (1882), pp. 301–329.

† Dickson, ibid., p. 16.

‡ One could have appealed to a well known result in quadratic form theory and assumed from the beginning that C was a prime.

Thus Ψ is equivalent to a binary form Ψ_1 with C' as the coefficient of Z^2. A transformation carrying Ψ into Ψ_1 carries F_1 into $F' = \cdots + C'Z^2$, and its contragedient transformation carrying f_1 into f', whose reciprocal is F', leaves the leading coefficient a_1 unaltered. Thus $a_1 = a'$, and f' is a form required in Theorem 1.

If the reciprocal F of f is improperly primitive, then f is properly primitive and, by virtue of the symmetry of the relation of reciprocity, Theorem 1 may be applied with f and F interchanged. We obtain thus the following result.

THEOREM 2. *If f is a properly primitive form with an improperly primitive reciprocal, then it is equivalent to a form f' such that $C' = 2\gamma$ and a' and γ are distinct odd primes not dividing $2\Omega\Delta$. Here, as before, a' is the leading coefficient of f' and C' the third coefficient of the reciprocal F' of f'.*

2. **Reduction to sum of squares.** We will assume now that the leading coefficient a of f and the third coefficient C of F are either distinct odd primes not dividing $\Omega\Delta$ or doubles of such primes.

Multiplying both members of (1) by a we obtain

$$af(x, y, z) = (ax + ty + sz)^2 + \Omega Cy^2 - 2\Omega Ryz + \Omega Bz^2, \tag{6}$$

where

$$ab - t^2 = \Omega C, \qquad ac - s^2 = \Omega B, \qquad ar - st = -\Omega R. \tag{7}$$

Next, multiplying both members of (6) by C and noting that by determinant theory

$$BC - R^2 = \Delta a, \tag{8}$$

we get

$$Caf(x, y, z) = C(ax + ty + sz)^2 + \Omega(Cy - Rz)^2 + \Omega\Delta az^2. \tag{9}$$

We now write

$$G(X, Y, Z) = CX^2 + \Omega Y^2 + \Omega\Delta aZ^2 \tag{10}$$

and speak of the so-constructed form G in the independent variables X, Y, Z as a form *associated* with f.

THEOREM 3. *Let f be a primitive ternary quadratic form, and let F be its reciprocal form. Employ the notation of* (1) *and* (2). *Let the leading coefficient a of f and the third coefficient C of F be either odd primes or doubles of such primes. Then if f represents an integer m, the associated form G in* (10) *represents aCm and in case $\Omega \equiv C \equiv 0 \pmod 2$ then*

$Y \equiv Z \pmod 2$. Conversely if the form G represents aCm and if $Y \equiv Z \pmod 2$ when $\Omega \equiv C \equiv 0 \pmod 2$, then f represents m.

*If the representation of aCm by G is proper, then that of m by f is likewise proper. The converse of the last statement holds with reservation; namely, if a proper representation x, y, z of m by f is given, then the related representation X, Y, Z of aCm by F is also proper if a and C are prime to m.**

The first part of the theorem is obvious in view of (9). We assume next that G represents aCm. There exist, therefore, integers X, Y, Z such that

$$aCm = CX^2 + \Omega Y^2 + \Omega\Delta a Z^2. \tag{11}$$

It follows from (11) that $\Omega Y^2 + \Omega\Delta a Z^2 \equiv 0 \pmod C$. But by (8), $\Delta a \equiv -R^2 \pmod C$. Thus $\Omega(Y^2 - R^2Z) \equiv 0 \pmod C$, and, multiplying through by -1 and factoring, we get $\Omega(RZ+Y)(RZ-Y) \equiv 0 \pmod C$. Since C is an odd prime not dividing Ω or a double of such a prime, since $Y \equiv Z \pmod 2$ if $C \equiv \Omega \equiv 0 \pmod 2$ by hypothesis, and since R is odd in the latter case in view of the choice of a, we have either $RZ + Y \equiv 0$ or $RZ - Y \equiv 0 \pmod C$. Thus, there exists an integer y such that either

$$RZ + Y = Cy \tag{12_1}$$

or

$$RZ - Y = Cy, \tag{12_2}$$

that is, such that

$$\pm Y = Cy - RZ. \tag{13}$$

Substituting $(Cy - RZ)^2$ for Y^2 in (11), we get, in view of (8),

$$\begin{aligned} aCm &= CX^2 + \Omega(Cy - RZ)^2 + \Omega(BC - R^2)Z^2 \\ &= CX^2 + \Omega C^2y^2 - 2\Omega CRyZ + \Omega CBZ^2, \end{aligned}$$

whence

$$am = X^2 + \Omega Cy^2 - 2\Omega RyZ + \Omega BZ^2. \tag{14}$$

From (14) it follows that

$$X^2 + \Omega Cy^2 - 2\Omega RyZ + \Omega BZ^2 \equiv 0 \pmod a, \tag{15}$$

* For any particular m this condition may be fulfilled (save possibly for a factor 2) by a suitable choice of a and C.

and since, in view of (7), $\Omega C \equiv -t^2$, $\Omega R \equiv ts$, $\Omega B \equiv -s^2 \pmod{a}$, we have

$$X^2 - t^2y^2 - 2tsyZ - s^2Z^2 \equiv 0 \pmod{a}, \quad X^2 - (ty + sZ)^2 \equiv 0 \pmod{a}.$$

Since a is an odd prime or a double of such a prime, either $X - ty - sZ \equiv 0$ or $-X - ty - sZ \equiv 0 \pmod{a}$. Thus, there exists an integer x such that either

$$X - ty - sZ = ax \tag{16_1}$$

or

$$-X - ty - sZ = ax, \tag{16_2}$$

that is, such that

$$\pm X = ax + ty + sZ. \tag{17}$$

Substituting $(ax+ty+sz)^2$ for X^2 in (14) we obtain, in view of (7),

$$am = (ax + ty + sZ)^2 + (ab - t^2)y^2 + 2(ar - st)yZ + (ac - s^2)Z^2,$$

whence

$$m = ax^2 + by^2 + cZ^2 + 2ryZ + 2sxZ + 2txy.$$

We now let

$$z = Z. \tag{18}$$

Then $f(x, y, z) = m$, and, since x and y determined by (12) and (16) are integers and Z is one by assumption, the integer m is represented by f.

From (13), (17) and (18) it follows that every common prime factor p of x, y, z divides X, Y, Z. Therefore if X, Y, Z are relatively prime, then so are x, y, z. The converse of the first statement is true for every prime p not dividing either a or C. For, if such a prime p divides X, Y, Z, then by (13) it divides y, by (17) it divides x, and by the choice of z in (18) it divides z. Thus if x, y, z are relatively prime, then the only possible common factors of X, Y, Z are divisors of a or C. But by virtue of (11) and the choice of a and C such common factors divide m. Hence the reservation of the converse of the second part of the theorem is sufficient.

3. **Application to universal forms.** Universal ternary quadratic forms were studied by Dickson,* who found the necessary and sufficient conditions that a form

$$\Phi = ex^2 + gy^2 + hz^2 \tag{19}$$

* Dickson, ibid., p. 21. Also this Bulletin, vol. 35 (1929), pp. 55–59.

should represent all integers, by Oppenheim* who studied forms with cross products and their equivalence to certain type forms, and by the present author,† who obtained conditions for universality of a ternary form in terms of its generic characters.

In this section we propose to show how by means of Theorem 3 and the normalization in §1, the criteria for universality for general forms with cross product terms may be deduced directly from the information available for forms of type (19).

The universality criteria in terms of generic characters run as follows.‡

THEOREM A. *Let f be a properly primitive, indefinite, classic,§ ternary, quadratic form with reciprocal F and determinant D. The necessary and sufficient conditions that f be universal are*

$$D = 2k + 1 \quad \text{or} \quad 2(2k + 1), \qquad \Omega = \pm 1 \tag{20}$$

where k is an integer, and, for every odd prime p dividing $\Delta = D$

$$(F \mid p) = (-\Omega \mid p). \tag{21}$$

THEOREM B. *Let f_1 be a primitive, indefinite, non-classic, ternary, quadratic form. Consider an improperly primitive form $f = 2f_1$. Let F be the reciprocal and Ω, Δ the invariants of f. Then f_1 is universal if and only if $f = 2f_1$ satisfies the following conditions:*

$$\Omega = \pm 1, \tag{22}$$

the characters of f are

$$(-1)^{(F-1)/2} = (-1)^{(\Omega+1)/2}, \qquad (F \mid p) = (-\Omega \mid p) \tag{23}$$

for every odd prime p dividing $\Delta = D$, and, in case 4 divides Δ,

$$(-1)^{(F^2-1)/8} = 1. \tag{24}$$

Noticing that conditions of Theorem A and relations (8) and (7_1) imply that the conditions of the above mentioned theorem of Dickson are fulfilled for the associated form G of f (see (10)), we may at once conclude that these conditions are sufficient.

To achieve the same for Theorem B, however, we need some addi-

* Oppenheim, Quarterly Journal of Mathematics (Oxford), vol. 1 (1930), pp. 179–185.

† Quarterly Journal of Mathematics, vol. 4 (1933), pp. 147–158.

‡ Ibid., pp. 147–148. There was a misprint in the statement of condition (1.51), our (23_1).

§ A classic form is one in which the coefficients of the cross products are all even. A form is non-classic if some or all of these coefficients are odd.

tional information about the form Φ in (19). We will proceed therefore to study forms of the type (19) and procure information permitting us not only to draw the conclusions of Theorem B, but also those of Theorem A. The procedure will be somewhat simplified by taking advantage of the fact that when the necessary conditions are imposed upon f the associated form G is of type (19) with $g=1$.

We first prove the following lemma.

LEMMA 1. *Consider a form*

$$\Phi(x, y, z) = ex^2 + y^2 + hz^2 = y^2 + \phi(x, z), \qquad eh = d, \tag{25}$$

such that

(26) $eh<0$ *and* $(e, h)=1$,

(27) $-h$ *and* $-e$ *are quadratic residues of* e *and* h *respectively.*

Then

$$\Phi \sim \Phi_1 = 2xy + y^2 - dz^2. \tag{28}$$

From (26) and (27), in view of a theorem due to Legendre* it follows that there exist integers λ, μ, ν relatively prime in pairs such that $\Phi(\lambda, \mu, \nu)=0$. Hence, by (25), $\phi(\lambda, \nu) = -\mu^2$, $(\lambda, \nu)=1$. Therefore, there exists a transformation carrying ϕ into an equivalent form

$$\phi(x, z) = (-\mu^2, M, N), \qquad M^2 + N\mu^2 = -eh = -d. \tag{29}$$

Henceforth the proof proceeds as in a similar lemma proved elsewhere.†

If we let $z=0$, $y=1$ or 2 in (28), we get

$$\Phi_1 = 2x + 1 \text{ or } 4(x + 1). \tag{30}$$

Next if $d\equiv 2 \pmod 4$, take $z=1$, $y=2$. Then $\Phi_1=4(x+1)+d$. If d is odd take $z=1$, $y=1$ and $x=2u$ or $2u+1$ according as $d=4h-1$ or $4h+1$. Then $\Phi_1=4(u-h)+2$. In both cases Φ_1 is equal to any $4l+2$ by choice of y and u respectively. We thus obtain a second lemma.

LEMMA 2. *If* d *in* (28) *is odd or a double of an odd integer, then the form* Φ_1 *represents all integers. Should* $d\equiv 0 \pmod 4$, Φ_1 *would not be universal but would still represent all multiples of* 4 *by* (30_2).

We are ready now for the proof of Theorems A and B.

It is not difficult to see that the conditions (20)–(21) and (22)–(24) are necessary.‡

* For exposition see Dickson, *Introduction to the Theory of Numbers*, chap. 8.

† Quarterly Journal of Mathematics, vol. 4 (1933), §3, p. 150.

‡ Ibid., §§ 4 and 7.

We modify slightly the following portion of the proof just referred to.

If $\Delta \equiv 4 \pmod 8$, the only residues of $\Omega\Delta az^2$ modulo 64 are 0, $\Omega\Delta a$, 32, where $\Omega\Delta a \equiv 8 \pmod{16}$. If $C \equiv -\Omega+4 \pmod 8$ holds, we show as before that $CX^2+\Omega Y^2 \not\equiv 8, 24, 32, 40, 56 \pmod{64}$. Then the only possible residues modulo 64 which are congruent to 0 (mod 8) are 0, 16, 48 and those obtained by adding to them 0, 32, and $\Omega\Delta a$. But in this manner we obtain only seven residues 0, 16, 48, 32, $\Omega\Delta a$, $16+\Omega\Delta a$, $48+\Omega\Delta a$ which are congruent to 0 (mod 8) out of the possible eight. Therefore, $C \equiv -\Omega \pmod 8$ and (24) holds.

The sufficiency of the conditions (20)–(21) and (22)–(24) follows at once from the two lemmas of this section, in view of Theorem 3 and Theorems 1 and 2.

If f is universal, so is $-f$, and if conditions (20)–(21) or (22)–(24) hold for one, they also hold for the other. One of these has $\Delta < 0$, $\Omega > 0$. We may assume then that $\Omega = +1$. Then (10) becomes

$$G(X, Y, Z) = CX^2 + Y^2 + \Delta aZ^2. \tag{31}$$

Thus G is of type (25). Since f is indefinite, we have (26_1). Condition (26_2) holds by the choice of a and C. Next, condition (27) holds in view of (8), and (27_2) holds by (7_1), (20_1)–(21) or (23)–(24), according as f is properly or improperly primitive. Therefore, by Lemma 1, $G \sim \Phi_1$. If f is properly primitive, then $d = \Delta aC$ is odd or double an odd integer and hence (Lemma 2) Φ_1 and thus G represents all integers and therefore all multiples of aC. Then, by Theorem 3, f represents all integers.

If f is improperly primitive, $d = \Delta aC \equiv 0 \pmod 4$. By Lemma 2, Φ_1 represents all multiples of 4. That is, Φ_1 and therefore G represents all even multiples of aC. Then, by Theorem 3, f represents all even integers, and hence $f/2$ is universal.

In closing, it may be of interest to mention a theorem first conjectured and proved by Dickson:*

THEOREM C. *Every universal, classic or non-classic, ternary quadratic form is a zero form.*

A proof of this theorem in a more general setting was given by Albert.†

ST. LOUIS UNIVERSITY

* L. E. Dickson, *Studies in the Theory of Numbers*, Chicago, 1930, p. 17. Also cf. Dickson, *Modern Elementary Theory of Numbers*, Chicago, 1939, chap. 8.

† A. A. Albert, this Bulletin, vol. 39 (1933), pp. 585–588.

ON NUMERICAL BOUNDS IN SCHOTTKY'S THEOREM*

RAPHAEL M. ROBINSON

According to Schottky's theorem, a function $f(z)$, which is regular and different from ± 1 for $|z|<1$, and for which $f(0)=a$, is bounded in absolute value for $|z|\leqq r$ by a number depending only on a and r. The asymptotic behaviour of the bound in Schottky's theorem for $a\to\infty$ has been studied by Ostrowski,† and numerical bounds have been given by Pfluger‡ and Ahlfors;§ but the numerical bounds which have been given are not asymptotically equal to the best bound.

Let $K(a, r)$ be the best bound possible in Schottky's theorem, in the form stated above. The purpose of this paper is to obtain, in as simple a manner as we can, estimates for $K(a, r)$ which are close enough to make it possible to derive from them the asymptotic formula

$$8K(a, r) \sim (8|a|)^{(1+r)/(1-r)}$$

for $a\to\infty$ and fixed r. *In fact we shall show that*

$$8K(a, r) - 10 < (8|a| + 10)^{(1+r)/(1-r)},$$

while

$$8K(a, r) + 10 > (8|a| - 10)^{(1+r)/(1-r)},$$

provided $8|a|-10>0$; *and from these inequalities the asymptotic formula evidently follows.*

We assume as known that it is possible to map a triangle with zero angles, and sides orthogonal to the unit circle, conformally on the upper half-plane, the mapping being continuous on the boundary; the vertices of the triangle may be taken into 1, -1, and ∞. (The mapping function is essentially an elliptic modular function.) The mapping function may be continued analytically throughout the unit circle by means of the Schwarz reflection principle, since this circle may be completely covered by triangles obtained from the original triangle by successive reflections on the sides. For later pur-

* Presented to the Society, April 15, 1939.

† *Asymptotische Abschätzung des absolutes Betrages einer Funktion, die die Werte* 0 *und* 1 *nicht annimmt*, Commentarii Mathematici Helvetici, vol. 5 (1933), pp. 55–87.

‡ *Über numerische Schranken im Schottky'schen Satz*, Commentarii Mathematici Helvetici, vol. 7 (1935), pp. 159–170.

§ *An extension of Schwarz's lemma*, Transactions of this Society, vol. 43 (1938), pp. 359–364.

poses, it is important to notice that if we start with a triangle containing the origin, a point of that triangle by successive reflections is always taken further from the origin, since it is always reflected from the exterior to the interior of a circle orthogonal to the unit circle.* Each triangle is mapped on an upper or lower half-plane, the entire circle $|z|<1$ being mapped on a Riemann surface with branch points of infinite order at 1, -1, and ∞. The mapping function is regular for $|z|<1$, and different from ± 1. The inverse function is many-valued, but has no singularities except at 1, -1, and ∞.

By means of a linear transformation of the unit circle into itself, we can insure that the mapping function has any desired value, different from ± 1, at the origin. Let $T(z)=T(z, a)$ be the function thus obtained, for which $T(0)=a$ and $T'(0)>0$. If we form the function $T(z, a)$ with $a=f(0)$, and let $T^{-1}(w)$ denote the inverse (many-valued) function, then the function $g(z)=T^{-1}(f(z))$ will be regular for $|z|<1$, since $f(z)$ is different from ± 1; hence $g(z)$ will be single-valued, as soon as a definite branch of $T^{-1}(w)$ is chosen, which we shall do by imposing the condition $g(0)=0$. Then $g(z)$ satisfies the conditions of Schwarz's lemma, so that $|g(z)| \leqq |z|$. Since $f(z)=T(g(z))$, this leads to the conclusion that, for any function $f(z)$ with $f(0)=a$, which is regular and different from ± 1 in the unit circle, we have

$$|f(z)| \leqq \max_{|z'|\leqq r} T(z', a) \quad \text{for} \quad |z| \leqq r.$$

This is Schottky's theorem; and in fact the bound is exact, since $T(z, a)$ is an admissible function. The expression on the right is then equal to $K(a, r)$. In estimating this maximum, z' may be restricted to the intersection of the triangle (of our set of triangles covering the unit circle) containing the origin, and the circle $|z'| \leqq r$. For the circle $|z'| \leqq r$ is covered by the reflections of this part of the triangle containing the origin (since a point of the triangle outside of the circle is never reflected into the circle), and at two corresponding points, the values of $T(z', a)$ are equal or conjugate.

For the computation, it is convenient to map the unit circle $|z| \leqq 1$ on an upper half-plane $\Im\zeta \geqq 0$. Since the bound to be computed is the same for any value of a as for its conjugate, we may suppose for convenience that $\Im a \geqq 0$, so that the triangle which contains the origin corresponds to an upper half-plane in the mapping $w=T(z, a)$. We may then suppose that this triangle goes into the triangle $-1 \leqq \Re\zeta \leqq 1$, $|\zeta| \geqq 1$, in such a way that the vertices $\zeta=1$, -1, ∞, correspond to $w=1$, -1, ∞, respectively. The function $T(z, a)$ then goes over into

* This fact is used in the last sentence of the following paragraph.

a function $\Lambda(\zeta)$ which maps the triangle $-1 \leqq \Re\zeta \leqq 1$, $|\zeta| \geqq 1$, on the upper half-plane $\Im w \geqq 0$, in such a way that 1, -1, and ∞, are fixed. Let $z=0$ go into the point $\zeta=b$; then $\Lambda(b)=T(0, a)=a$. The circle $|z| \leqq r$ goes over into a circle whose highest point is at a distance $\gamma\Im b$ from the real axis, where γ is an abbreviation for $(1+r)/(1-r)$. The value of $K(a, r)$ is then less than or equal to the maximum of $|\Lambda(\zeta)|$ for ζ in the triangle and $\Im\zeta \leqq \gamma\Im b$, and this maximum is attained on the upper boundary $\Im\zeta=\gamma\Im b$, since $|\Lambda(\zeta)|$ increases upward on the vertical sides. On the other hand, the region over which we had to compute the maximum of $|\Lambda(\zeta)|$ contains a point of $\Im\zeta=\gamma\Im b$. Hence $K(a, r)$ is included between the minimum and maximum of $|\Lambda(\zeta)|$ for ζ in the triangle and $\Im\zeta=\gamma\Im b$. (That the problem can be reduced to this has been recognized by others; but they seem to have overlooked the simple method of calculating a bound given below.)

Let $\phi(\zeta)$ be the function which maps our triangle in the ζ-plane onto the region $\Im w \geqq 0$, $|w| \geqq 1$, in such a way that 1, -1, and ∞ are fixed. Then $\Lambda(\zeta)=[\phi(\zeta)+1/\phi(\zeta)]/2$. We note further that $\Lambda(\zeta)$ maps the arc of the circle $|\zeta+1|=2$ extending from 1 to $-1+2i$ (which may be characterized as a circular arc through the vertex 1 which is orthogonal to the opposite side of the triangle as well as to the real axis) onto an arc of the same circle extending from 1 to -3. For if we transform the mapping by taking the upper half-plane into the unit circle by a linear transformation, in such a way that the vertices of the triangle go into points which are equally spaced around the unit circle, the corresponding statement will be evident from the symmetry of the figure. From this we conclude that $\Lambda(\zeta)$ maps the region $-1 \leqq \Re\zeta \leqq 1$, $|\zeta+1| \geqq 2$, onto the region $\Im w \geqq 0$, $|w+1| \geqq 2$, the points 1 and ∞ being fixed, and $-1+2i$ going into -3. That is the same as saying that $\Lambda(2\zeta-1)$ takes the region $0 \leqq \Re\zeta \leqq 1$, $|\zeta| \geqq 1$, into the same region in the w-plane, the points 1 and ∞ being fixed, and i going into -3. But $2\phi(\zeta)^2-1$ does exactly the same, hence $\Lambda(2\zeta-1) = 2\phi(\zeta)^2-1$. Thus both $\Lambda(\zeta)$ and $\Lambda(2\zeta-1)$ are expressed in terms of $\phi(\zeta)$, and hence brought into relation to each other.

By means of the function $t=ie^{\pi\zeta/2i}$ the half-strip $-1 \leqq \Re\zeta \leqq 1, \Im\zeta \geqq 0$ is taken into the region $\Im t \geqq 0$, $|t| \geqq 1$. This transformation takes our triangle in the ζ-plane into a region in the upper half of the t-plane, bounded by the part of the real axis to the left of -1 and the part to the right of 1, and by a curve C joining -1 and 1, whose distance from the origin is always between 1 and $e^{\pi/2}$. We may regard $\Lambda(\zeta)$ as a function of t, for t in the part of the upper half-plane outside of C, and continue across the real axis by the Schwarz reflection principle, so that $\Lambda(\zeta)$ is a regular function of t, for t outside of C and its reflec-

tion in the real axis, except for a simple pole at ∞.* It is clear from the symmetry that the expansion of $\Lambda(\zeta)$ in the neighborhood of $t=\infty$ is of the form $\Lambda(\zeta)=ct+c_1/t+c_3/t^3+\cdots$, all the coefficients being real, and the first coefficient c being positive. From the functional relation obtained above, the coefficients c, c_1, c_3, $\cdots$ could be calculated. However, we need only the value of c. Now if we replace ζ by $2\zeta-1$, then t is replaced by t^2, so that besides the relation $\Lambda(\zeta)\sim ct$, we have also $\Lambda(2\zeta-1)\sim ct^2$. But from the equations expressing $\Lambda(\zeta)$ and $\Lambda(2\zeta-1)$ in terms of $\phi(\zeta)$, we have $\Lambda(\zeta)\sim\phi(\zeta)/2$ and $\Lambda(2\zeta-1)\sim 2\phi(\zeta)^2$. Eliminating the $\phi(\zeta)$ leads to the conclusion that $ct^2\sim 8c^2t^2$, from which we find $c=1/8$. From this we see that $8\Lambda(\zeta)-t\to 0$ for $t\to\infty$. Thus the maximum of $|8\Lambda(\zeta)-t|$ will occur for t on the curve C. Since the curve is symmetric with respect to the imaginary axis, the part in the first quadrant is taken into the segment from 0 to 8 by the function $8\Lambda(\zeta)$. It is clear then that $|8\Lambda(\zeta)-t|\leqq(64+e^\pi)^{1/2}<10$ for t on C and hence for all values of t considered. Now $|t|=e^{(\pi/2)\Im\zeta}$, so that

$$8|\Lambda(\zeta)|-10<e^{(\pi/2)\Im\zeta}<8|\Lambda(\zeta)|+10$$

for all points ζ of our triangle. Hence for $\Im\zeta=\gamma\Im b$, we have

$$8|\Lambda(\zeta)|-10<e^{(\pi/2)\Im\zeta}=[e^{(\pi/2)\Im b}]^\gamma<[8|\Lambda(b)|+10]^\gamma,$$

and, provided $8|\Lambda(b)|-10>0$,

$$8|\Lambda(\zeta)|+10>e^{(\pi/2)\Im\zeta}=[e^{(\pi/2)\Im b}]^\gamma>[8|\Lambda(b)|-10]^\gamma.$$

Remembering that $\Lambda(b)=a$, and that $K(a,r)$ is between the maximum and minimum of $|\Lambda(\zeta)|$ for $\Im\zeta=\gamma\Im b$, we see that these inequalities yield the desired bounds for $K(a,r)$.

UNIVERSITY OF CALIFORNIA

* The function $\Lambda(\zeta)$ is evidently a single-valued function of t in the region mentioned, since in continuing across either segment of the real axis we obtain its value at a point t of the lower half-plane as the conjugate of its value at $\bar{t}$.

ON SERRET'S INTEGRAL FORMULA*

G. E. RAYNOR

In 1901 F. Morley† evaluated the integral

$$\int_0^{\pi/2} \phi^n \log^m (2\cos\phi)\,d\phi \tag{1}$$

for even nonnegative integral values of n, and m a positive integer. His method was that of contour integration. However, the results could have been obtained with the help of a formula derived by A. R. Forsyth a few years earlier‡ and also from a formula derived by Cauchy as early as 1825.§ Cauchy's formula is

$$\int_0^{\pi/2} \cos^p\phi \cos q\phi\, d\phi = \frac{\pi}{2^{p+1}} \frac{\Gamma(p+1)}{\Gamma((p+q)/2+1)\Gamma((p-q)/2+1)}, \qquad R(p) > -1.$$

Here p and q are in general complex numbers with p subject to the restriction indicated. If this relation be written in the form

$$\int_0^{\pi/2} (2\cos\phi)^p \cos q\phi\, d\phi = \frac{\pi}{2} \frac{\Gamma(p+1)}{\Gamma((p+q)/2+1)\Gamma((p-q)/2+1)}$$

and then differentiated m times with respect to p and n times (n even) with respect to q, and then p and q set equal to zero, the values of integrals of the form (1) are obtained quite easily.

In 1843 Serret‖ obtained the formula

$$\int_0^{\pi/2} \cos^p\phi \sin q\phi\, d\phi = \frac{\Gamma(p+1)}{\Gamma((p+q)/2+1)\Gamma((p-q)/2+1)} \int_0^1 \frac{t^{(p-q)/2} - t^{(p+q)/2}}{(1+t)^{p+1}(1-t)}\,dt, \qquad R(p) > -1. \tag{2}$$

* Presented to the Society, October 28, 1939.

† This Bulletin, vol. 7 (1901), p. 390.

‡ A. R. Forsyth, *Evaluation of two definite integrals*, Quarterly Journal of Mathematics, vol. 27 (1895), p. 221.

§ A. Cauchy, *Mémoire sur les Intégrales Définies Prises entre des Limites Imaginaires*, Paris, 1825, p. 40.

‖ Alfred Serret, Journal de Mathématiques Pures et Appliquées, vol. 8 (1845), p. 7. The formula as given in Serret's article contains two misprints which are repeated in Encyklopädie der mathematischen Wissenschaften, vol. 2, part 2, no. A12, p. 853. They are corrected below.

This result has been quite neglected in the literature on definite integrals. It is the purpose of this note to show how this formula, together with certain well known results of gamma function theory, may be used to study an important class of definite integrals and incidentally to *supplement* certain recent results.

In the following we have employed a number of improper integrals and have also repeatedly differentiated under the integral sign. It is, of course, necessary to make sure that all the integrals used exist, and to justify the double limit processes. This the writer has done, but the details have been omitted in the belief that the reader can easily supply them for himself.

In two interesting papers* Rutledge and Douglass have treated a number of definite integrals which are related to those of the type (1) with n *odd* and to the integral $\int_0^1 [(\log u)/u] \log^2 (1+u)du$ appearing in the title of their first paper. Two of their principal results are embodied in the following formulas:

$$\int_0^{\pi} \phi \log^2 \left(2 \cos \frac{\phi}{2}\right) d\phi = A_4 + \frac{31\pi^4}{480}, \tag{3}$$

$$\int_0^1 \frac{\log u}{u} \log^2 (1+u)du = A_4 - \frac{\pi^4}{288}. \tag{4}$$

The constant A_4 is defined by the equation

$$A_4 = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{(n+1)^2}\left(\frac{1}{1^2} + \frac{1}{2^2} + \cdots + \frac{1}{n^2}\right). \tag{5}$$

By an obvious transformation equation (3) is seen to be equivalent to

$$\int_0^{\pi/2} \phi \log^2 (2 \cos \phi)d\phi = \frac{A_4}{4} + \frac{31\pi^4}{1920}. \tag{6}$$

With the help of expansions in series, contour integration and algebraic combinations various integrals of the nature of that in (4) have been evaluated.† We wish to show how Serret's formula may be used to supplement this work and, as an example, shall evaluate an integral algebraically independent of those given in the table of integrals of order four by Rutledge and Douglass.†

* 1. George Rutledge and R. D. Douglass, *Evaluation of* $\int_0^1 [(\log u)/u] \log^2 (1+u)du$ *and related definite integrals*, American Mathematical Monthly, vol. 41 (1934), pp. 29–36. 2. *Table of definite integrals*, American Mathematical Monthly, vol. 45 (1938), pp. 525–530.

† G. Rutledge and R. D. Douglass, loc. cit. (2), p. 527.

We set down here some formulas needed later and which are to be found in Nielsen's book on the gamma function:*

$$\Psi(1+x) = -s_1 + s_2 x - s_3 x^2 + s_4 x^3 - s_5 x^4 + \cdots; \tag{7}$$

$$\beta(1+x) = \sigma_1 - \sigma_2 x + \sigma_3 x^2 - \sigma_4 x^3 + \sigma_5 x^4 - \cdots; \tag{8}$$

$$s_1 = C = \text{Euler's constant}; \qquad s_n = \sum_{p=1}^{\infty} \frac{1}{p^n}, \qquad n > 1; \tag{9}$$

$$\sigma_n = \sum_{p=1}^{\infty} \frac{(-1)^{p-1}}{p^n}, \qquad n \geqq 1; \tag{10}$$

$$\sigma_1 = \log 2, \qquad \sigma_n = (1 - 1/2^{n-1})s_n, \qquad n > 1; \tag{11}$$

$$s_{2n} = \frac{B_n(2\pi)^{2n}}{(2n)!2}, \qquad B_1 = \frac{1}{6}, \qquad B_2 = \frac{1}{30}; \tag{12}$$

$$\beta(x) = \int_0^1 \frac{t^{x-1}}{1+t}\,dt, \qquad R(x) > 0; \tag{13}$$

$$\Psi(x) + C = \int_0^1 \frac{1 - t^{x-1}}{1-t}\,dt, \qquad R(x) > 0; \tag{14}$$

$$\eta(x) = \int_0^1 \frac{\log 2 - \log(1+t)}{1-t}\, t^{x-1}dt, \qquad R(x) > 0; \tag{15}$$

$$\beta^2(x) = \Psi^{(1)}(x) - 2\eta'(x). \tag{16}$$

The notation $\Psi^{(n)}(x)$ means the nth derivative of $\Psi(x)$.

Now let each side of (2) be differentiated with respect to q and then q set equal to zero. We obtain

$$\int_0^{\pi/2} \phi \cos^p \phi\, d\phi = -\frac{\Gamma(p+1)}{\Gamma^2(p/2+1)} \int_0^1 \frac{t^{p/2}}{(1+t)^{p+1}(1-t)} \log t\, dt. \tag{17}$$

Set

$$J(p) = \Gamma(p+1)/\Gamma^2(p/2+1). \tag{18}$$

Then since, by definition, $\Psi(x) = \Gamma^{(1)}(x)/\Gamma(x)$, it follows with the help of (7), that

$$J^{(1)}(0) = 0, \qquad J^{(2)}(0) = s_2/2. \tag{19}$$

Also let

$$T(p) = (t^{1/2}/(1+t))^p. \tag{20}$$

Then

* N. Nielsen, *Handbuch der Theorie der Gammafunktion*, equations (7), (8), (9), (10), (11), p. 37; (12), p. 46; (13), p. 169; (14), p. 170; (15), p. 195; (16), p. 194.

$$T^{(n)}(p) = (t^{1/2}/(1+t))^p \log^n t^{1/2}/(1+t) = (t^{1/2}/(1+t))^p[(\log t)/2 - \log(1+t)]^n. \tag{21}$$

Now differentiating (17) twice with respect to p and then letting $p=0$, we have by (19) and (21)

$$\int_0^{\pi/2} \phi \log^2 \cos\phi d\phi = -\frac{1}{2} s_2 \int_0^1 \frac{\log t dt}{(1+t)(1-t)} - \int_0^1 \frac{[(\log t)/2 - \log(1+t)]^2 \log t}{(1+t)(1-t)} dt. \tag{22}$$

From (17), for $p=0$,

$$\int_0^{\pi/2} \phi d\phi = -\int_0^1 \frac{\log t dt}{(1+t)(1-t)},$$

or

$$\frac{\pi^2}{8} = -\int_0^1 \frac{\log t dt}{(1+t)(1-t)}. \tag{23}$$

Hence (22) becomes, since by (12) $s_2=\pi^2/6$ and $s_4=\pi^4/90$,

$$\begin{aligned}
&\int_0^{\pi/2} \phi \log^2 \cos\phi d\phi \\
&= \frac{15}{16} s_4 - \int_0^1 \frac{[(\log t)/2 - \log(1+t)]^2 \log t}{(1+t)(1-t)} dt \\
&= \frac{15}{16} s_4 - \frac{1}{8}\int_0^1 \frac{\log^3 t}{1+t} dt - \frac{1}{8}\int_0^1 \frac{\log^2 t}{1-t} dt \\
&\quad + \frac{1}{2}\int_0^1 \frac{\log(1+t)\log^2 t}{1+t} dt + \frac{1}{2}\int_0^1 \frac{\log(1+t)\log^2 t}{1-t} dt \\
&\quad - \frac{1}{2}\int_0^1 \frac{\log^2(1+t)\log t}{1+t} dt - \frac{1}{2}\int_0^1 \frac{\log^2(1+t)\log t}{1-t} dt.
\end{aligned} \tag{24}$$

Now from (13), (8) and (11) we have

$$\int_0^1 \frac{\log^3 t}{1+t} dt = \beta^{(3)}(1) = -6\sigma_4 = -\frac{21}{4} s_4. \tag{25}$$

From (14) and (7) it follows that

$$\int_0^1 \frac{\log^3 t}{1-t} dt = -\Psi^{(3)}(1) = -6s_4. \tag{26}$$

Also from (4) we obtain, by an integration by parts,

$$\int_0^1 \frac{\log(1+t)\log^2 t}{1+t}\,dt = -A_4 + \frac{5}{16}s_4. \tag{27}$$

To obtain the 4th integral on the right of (24) we first modify (15). The integral (15) cannot be written as the difference of the two integrals

$$\int_0^1 \frac{\log 2}{1-t}t^{x-1}dt, \qquad \int_0^1 \frac{\log(1+t)}{1-t}t^{x-1}dt$$

since they do not separately exist. However, if we differentiate each side of (15) with respect to x, we may write

$$\eta^{(1)}(x) = \log 2\int_0^1 \frac{\log t}{1-t}t^{x-1}dt - \int_0^1 \frac{\log(1+t)\log t}{1-t}t^{x-1}dt,\ R(x) > 0,$$

which by (14) becomes

$$\eta^{(1)}(x) = -\sigma_1\Psi^{(1)}(x) - \int_0^1 \frac{\log(1+t)\log t}{1-t}t^{x-1}dt. \tag{28}$$

This, with the help of (16), gives

$$\int_0^1 \frac{\log(1+t)\log t}{1-t}t^{x-1}dt = -\sigma_1\Psi^{(1)}(x) - (1/2)\Psi^{(2)}(x) + \beta(x)\beta^{(1)}(x).$$

Hence finally,

$$\begin{aligned}\int_0^1 \frac{(\log 1+t)\log^2 t}{1-t}\,dt &= -\sigma_1\Psi^{(2)}(1) - (1/2)\,\Psi^{(3)}(1) + (\beta^{(1)}(1))^2 + \beta(1)\beta^{(2)}(1)\\ &= 2\sigma_1 s_3 - 3s_4 + \sigma_2^2 + 2\sigma_1\sigma_3 = 7\sigma_1 s_3/2 - 19s_4/8.\end{aligned} \tag{29}$$

From the 5th integral on the right of (24) we obtain, by an integration by parts, the relation

$$\int_0^1 \frac{\log^2(1+t)\log t}{1+t}\,dt = -\frac{1}{3}\int_0^1 \frac{\log^3(1+t)}{t}\,dt.$$

The value of the last integral is given by formula (7) of the table referred to above.* Thus,

* Rutledge and Douglass, *Table of definite integrals*, loc. cit.

$$\int_0^1 \frac{\log^2(1+t)\log t}{1+t}\,dt = -\frac{1}{2}A_4 + \frac{1}{32}s_4. \tag{30}$$

Hence, substituting the results of (25), (26), (27), (29) and (30) in (24) we have the relation

$$\int_0^{\pi/2} \phi \log^2 \cos\phi d\phi = \frac{83}{64}s_4 + \frac{7}{4}\sigma_1 s_3 - \frac{1}{4}A_4 - \frac{1}{2}\int_0^1 \frac{\log^2(1+t)\log t}{1-t}\,dt. \tag{31}$$

Now, from (6) by writing $\log(2\cos\phi) = \log 2 + \log\cos\phi$ we have, after a few reductions, the relation

$$\int_0^{\pi/2} \phi \log^2 \cos\phi d\phi = -2\sigma_1 \int_0^{\pi/2} \phi \log\cos\phi d\phi - \frac{3}{4}\sigma_1^2 s_2 + \frac{93}{64}s_4 + \frac{1}{4}A_4. \tag{32}$$

The integral on the right of (32) may be evaluated completely by Serret's formula by differentiating the two sides of (17) once with respect to p, setting $p=0$, and then evaluating the resulting integrals in a manner analogous to that of the last paragraph. In this way it is found that

$$\int_0^{\pi/2} \phi \log\cos\phi d\phi = -\frac{7}{16}s_3 - \frac{3}{4}\sigma_1 s_2. \tag{33}$$

Then by (33) it is found that (32) becomes

$$\int_0^{\pi/2} \phi \log^2 \cos\phi d\phi = \frac{93}{64}s_4 + \frac{7}{8}\sigma_1 s_3 + \frac{3}{4}\sigma_1^2 s_2 + \frac{1}{4}A_4. \tag{34}$$

If now the right sides of (31) and (34) be equated, we get

$$\int_0^1 \frac{\log^2(1+t)\log t}{1-t}\,dt = -\frac{5}{16}s_4 + \frac{7}{4}\sigma_1 s_3 - \frac{3}{2}\sigma_1^2 s_2 - A_4. \tag{35}$$

This result, together with some of those above, serves to show that Serret's formula can be made a useful tool in studying certain types of definite integrals. Various other results may be obtained, both in connection with integrals of the orders treated here and of higher orders. However, we leave such considerations for treatment on another occasion.

The nature of the constant A_4 here remains undetermined just as in the papers of Rutledge and Douglass. Whether or not it can be rationally expressed in terms of the constants s_1, σ_1, s_3 and π is an open question. Some light may be thrown on the problem by a further study of the function $\xi_1(x)$ treated briefly by Nielsen.* His definition is as follows,

$$\xi_1(x) = \int_0^1 \frac{\log(1+t)}{1+t} t^{x-1}dt, \qquad R(x) > 0. \tag{36}$$

From this equation and (27) it follows that

$$A_4 = 5s_4/16 - \xi_1^{(2)}(1). \tag{37}$$

This in itself, of course, sheds no light but if a relation analogous to (16) could be found involving the function $\xi_1(x)$, it would seem that the question could be answered.

LEHIGH UNIVERSITY

THE COMPUTATION OF THE SMALLER COEFFICIENTS OF $J(\tau)$

HERBERT S. ZUCKERMAN†

The purpose of this note is to call attention to the fact that the first twenty-five coefficients $a_0, a_1, \cdots, a_{24}$ in the expansion

$$1728J(\tau) = e^{-2\pi i\tau} + \sum_{n=0}^{\infty} a_n e^{2\pi i n\tau} \tag{1}$$

can be computed with relative ease, making use of H. Gupta's tables‡ of the partition function which extend to $n=600$.

From the multiplicator equation§ of fifth order of $J(\tau)$ we have

$$1728J(\tau) = y^{-1} + 6\cdot 5^3 + 63\cdot 5^5 y + 52\cdot 5^8 y^2 + 63\cdot 5^{10} y^3 + 6\cdot 5^{13} y^4 + 5^{15} y^5, \tag{2}$$

with

* N. Nielsen, loc. cit., p. 233.

† Harrison Research Fellow.

‡ *A table of partitions*, Proceedings of the London Mathematical Society, vol. 39 (1935), pp. 142–149; *A table of partitions* II, Proceedings of the London Mathematical Society, vol. 42 (1937), pp. 546–549.

§ Klein-Fricke, *Vorlesungen über die Theorie der elliptischen Modulfunktionen*, vol. 2, p. 61, formula (11), with the values given in vol. 2, p. 64, (5) and vol. 1, p. 154, (1).

$$y = 5^{-3}\frac{\Delta(\tau, 1/5)^{1/4}}{\Delta(\tau, 1)^{1/4}} = e^{2\pi i\tau}\phi(e^{10\pi i\tau})^6\phi(e^{2\pi i\tau})^{-6}, \quad \phi(x) = \prod_{m=1}^{\infty}(1 - x^m).$$

On the other hand we have*

$$(3)\quad \sum_{n=0}^{\infty} p(25n + 24)e^{2\pi i n\tau} = \phi(e^{2\pi i\tau})^{-1}e^{-2\pi i\tau}\{63\cdot 5^2 y + 52\cdot 5^5 y^2 + 63\cdot 5^7 y^3 + 6\cdot 5^{10} y^4 + 5^{12} y^5\}.$$

Combining (2) and (3) we find

$$1728J(\tau) = y^{-1} + 6\cdot 5^3 + 5^3\phi(e^{2\pi i\tau})e^{2\pi i\tau}\sum_{n=0}^{\infty} p(25n + 24)e^{2\pi i n\tau}$$

and hence

$$(4)\quad x^{-1} + \sum_{n=0}^{\infty} a_n x^n = x^{-1}\phi(x^5)^{-6}\phi(x)^6 + 6\cdot 5^3 + 5^3 x\phi(x)\sum_{n=0}^{\infty} p(25n + 24)x^n.$$

Equation (4) may be used to compute the a_n. The first term of the right member can be expanded with the aid of Jacobi's formula for $\phi(x)^3$ while the expansion of $\phi(x)$ in the second term is given by Euler's formula. Thus we may write (4) as

$$(5)\quad x^{-1} + \sum_{n=0}^{\infty} a_n x^n = \sum_{n=-1}^{\infty} b_n x^n + 5^3(x - x^2 - x^3 + x^6 + x^8 - x^{13} - x^{16} + x^{23} + x^{27} - \cdots)\cdot\sum_{n=0}^{\infty} p(25n + 24)x^n,$$

with the values

$b_{-1}=1$,	$b_0=744$,	$b_1=9$,	$b_2=10$,
$b_3=-30$,	$b_4=6$,	$b_5=-25$,	$b_6=96$,
$b_7=60$,	$b_8=-250$,	$b_9=45$,	$b_{10}=-150$,
$b_{11}=544$,	$b_{12}=360$,	$b_{13}=-1230$,	$b_{14}=184$,
$b_{15}=-675$,	$b_{16}=2310$,	$b_{17}=1410$,	$b_{18}=-4830$,
$b_{19}=750$,	$b_{20}=-2450$,	$b_{21}=8196$,	$b_{22}=4920$,
$b_{23}=-16180$,	$b_{24}=2376$, $\cdots$.		

* G. N. Watson, *Ramanujans Vermutung über Zerfällungsanzahlen*, Journal für die reine und angwandte Mathematik, vol. 179 (1938), pp. 97–128; also H. S. Zuckerman, *Identities analogous to Ramanujan's identities involving the partition function*, Duke Mathematical Journal, vol. 5 (1939), pp. 88–110.

By using (5) and the table of partitions, the values of the a_n may be found by mere additions, subtractions, and a single multiplication by 5^3.

The following list contains the values of the a_n computed with the aid of Gupta's tables. The values for $n \leqq 7$ agree with those given by W. E. H. Berwick.*

$$
\begin{aligned}
a_0 &= 744,\\
a_1 &= 1\ 96884,\\
a_2 &= 214\ 93760,\\
a_3 &= 8642\ 99970,\\
a_4 &= 2\ 02458\ 56256,\\
a_5 &= 33\ 32026\ 40600,\\
a_6 &= 425\ 20233\ 00096,\\
a_7 &= 4465\ 69940\ 71935,\\
a_8 &= 40149\ 08866\ 56000,\\
a_9 &= 3\ 17644\ 02297\ 84420,\\
a_{10} &= 22\ 56739\ 33095\ 93600,\\
a_{11} &= 146\ 21191\ 14995\ 19294,\\
a_{12} &= 874\ 31371\ 96857\ 75360,\\
a_{13} &= 4872\ 01011\ 17981\ 42520,\\
a_{14} &= 25497\ 82738\ 94105\ 25184,\\
a_{15} &= 1\ 26142\ 91646\ 57818\ 43075,\\
a_{16} &= 5\ 93121\ 77242\ 14450\ 58560,\\
a_{17} &= 26\ 62842\ 41315\ 07752\ 45160,\\
a_{18} &= 114\ 59912\ 78844\ 47865\ 13920,\\
a_{19} &= 474\ 38786\ 80123\ 41688\ 13250,\\
a_{20} &= 1894\ 49976\ 24889\ 33900\ 28800,\\
a_{21} &= 7318\ 11377\ 31813\ 75192\ 45696,\\
a_{22} &= 27406\ 30712\ 51362\ 46549\ 29920,\\
a_{23} &= 99710\ 41659\ 93718\ 26935\ 33820,\\
a_{24} &= 3\ 53074\ 53186\ 56142\ 70998\ 77376.
\end{aligned}
$$

UNIVERSITY OF PENNSYLVANIA

* *An invariant modular equation of the fifth order*, Quarterly Journal of Mathematics, vol. 47 (1916), pp. 94–103.

NOTE ON A PRINCIPAL AXIS TRANSFORMATION FOR NON-HERMITIAN MATRICES

JOHN WILLIAMSON

In a recent note the following theorem was proved.†

THEOREM 1. *If A is a matrix, over the complex field, of r rows and s columns, there exist two unitary matrices U and V, such that*

$$UAV = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix},$$

where D is a real diagonal matrix $[d_1, d_2, \cdots, d_k]$ with positive elements d_i.

For completeness the following result may be added: *The elements d_i are determined uniquely as the positive square roots of the nonzero characteristic roots of the positive hermitian matrix AA^*, where A^* is the conjugate transposed of A.*

The elements d_i thus form a complete set of invariants for the matrix A under such unitary transformations. This, together with the theorem itself, may be proved as follows. If $r \leqq s$ and k is the rank of A, $A = (P, 0)V$, where V is unitary and P is the positive hermitian matrix of order r and rank k uniquely determined‡ by the equation $AA^* = P^2$. The hermitian matrix P is unitarily equivalent to the diagonal matrix of order r, whose first k elements are d_i, $(i = 1, 2, \cdots, k)$. In case $r \geqq s$, the polar representation

$$A = U\begin{pmatrix} P \\ 0 \end{pmatrix}$$

may be used, to prove the desired result.

The following theorem may also be of interest.

THEOREM 2. *Let A and B be two matrices, over the complex field, of r rows and s columns. Necessary and sufficient conditions that there exist two unitary matrices U and V, such that*

$$UAV = A_1, \qquad UBV = B_1, \tag{1}$$

where A_1 and B_1 are diagonal matrices, are that

† Carl Eckart and Gale Young, *A principal axis transformation for non-hermitian matrices*, this Bulletin, vol. 52 (1939), pp. 118–121.

‡ John Williamson, *A polar representation of singular matrices*, this Bulletin, vol. 41 (1935), pp. 118–123.

$$AB^* = f(BA^*), \qquad B^*A = f(A^*B), \tag{2}$$

where $f(x)$ is a polynomial in x.

We first prove the necessity of the conditions. Since $A_1B_1^*$ is a diagonal matrix, $A_1B_1^*$ is normal† and therefore $A_1B_1^* = f(B_1A_1^*)$. Since A_1 and B_1^* are both diagonal matrices, $A_1B_1^*$ coincides with $B_1^*A_1$ except for zero elements and so does $B_1A_1^*$ with $A_1^*B_1$. Therefore $B_1^*A_1 = f(A_1^*B_1)$ and, as a consequence of (1), (2) is true.

To prove the sufficiency we note that there is no loss in generality in assuming that A is of the form

$$\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix},$$

where D is a real nonsingular diagonal matrix. Let

$$\begin{pmatrix} H & L \\ M & K \end{pmatrix}$$

be a partition of B similar to that of A. Since AB^* is normal, $AB^*BA = BAAB^*$ and by a simple calculation we find that

$$MDDM^* = (MD)(MD)^* = 0.$$

Hence $MD = 0$ and, since D is nonsingular, $M = 0$. In a similar manner by using the fact that B^*A is normal we find that L is zero. Conditions (2) now reduce to

$$DH^* = f(HD), \qquad H^*D = f(DH). \tag{3}$$

Since $f(HD) = D^{-1}f(DH)D$, it follows from (3) that $D^2H^* = H^*D^2$. But D is a positive definite real matrix. Hence $DH^* = H^*D$ and consequently $DH = HD$. Since DH^* is normal, H^* and, therefore H, is normal. Since H is normal‡ and commutative with D, there exists a unitary matrix U_1, such that $U_1DU_1^* = D$ and $U_1HU_1^* = D_1$, where D_1 is diagonal but not necessarily real. By Theorem 1 there exist two unitary matrices U_2 and V_2, such that $U_2KV_2 = D_2$, where D_2 is diagonal and real. If

$$U = \begin{pmatrix} U_1 & 0 \\ 0 & U_2 \end{pmatrix}, \qquad V = \begin{pmatrix} U_1^* & 0 \\ 0 & V_2 \end{pmatrix},$$

then

† Aurel Wintner, *Spektraltheorie der unendlichen Matrizen*, Leipzig, 1929, p. 24; John Williamson, *Matrices normal with respect to an hermitian matrix*, American Journal of Mathematics, vol. 60 (1938), p. 355.

‡ Aurel Wintner, op. cit., p. 24.

$$UAV = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}, \qquad UBV = \begin{pmatrix} D_1 & 0 \\ 0 & D_2 \end{pmatrix}.$$

Since the elements of D and D_2 are real, we have the following corollary.

COROLLARY 1. *If k is the smaller of the ranks of A and B, A_1 and B_1 may be so determined that at most k of the elements of A_1 and B_1 are not real.*

If $f(x)=x$, AB^* and B^*A are both hermitian and D_1 must be real, so that B_1 is real. Conversely, if B_1 is real, $A_1B_1^*$ and $B_1^*A_1$ are both hermitian and therefore AB^* and B^*A are both hermitian. This result gives Theorem 2 of the note quoted in the first footnote. If $f(x)=1/x$, AB^* and B^*A are both unitary and the elements of D_1 are $d_i^{-1}e^{i\theta_i}$.

When $r=s$, so that the matrices A and B are square, condition (2) may be replaced by the condition that *AB^* be normal and unitarily equivalent to B^*A*. For AB^* is similar to B^*A and, since both are normal, AB^* is unitarily equivalent to B^*A.

Both Theorems 1 and 2 are true in the real field, if A^* denotes the transposed of A and "unitary" is replaced by "orthogonal," except that D_1 may have two-rowed matrices of the form

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

in its diagonal.† In the simplest case, when AB^* and B^*A are both symmetric, D_1 will necessarily be diagonal, and Theorem 2 of the paper, quoted in the first footnote, is true in the real field, if "unitary" is replaced by "orthogonal" and "hermitian" by "symmetric."

THE JOHNS HOPKINS UNIVERSITY

† F. D. Murnaghan and Aurel Wintner, *A canonical form for real matrices under orthogonal transformations*, Proceedings of the National Academy of Sciences, vol. 17 (1931), pp. 417–420.

ON THE BASIS THEOREM FOR INFINITE SYSTEMS OF DIFFERENTIAL POLYNOMIALS*

E. R. KOLCHIN

Introduction. Let $\mathcal{F}$ be a differential field of characteristic zero.† We consider an infinite system Σ of differential polynomials in the letters $y_1, \cdots, y_n$, the coefficients of the differential polynomials being in $\mathcal{F}$.‡

A finite set Φ of forms in Σ is called a *basis* of Σ if, for every form G in Σ, there is a positive integer p, dependent on G, such that G^p is in the differential ideal of Φ. If a single p will serve for every G in Σ, then we shall call the basis *strong*.

It has been shown that every system has a basis.§ Raudenbush has shown further,‖ that there exist systems, not every basis of which is strong. It is now natural to ask whether or not every system of forms contains at least one strong basis.

We answer this question in the negative by showing that even a perfect differential ideal of forms may have no strong basis. The perfect differential ideal with which we work is the one generated by the form uv in the two unknowns u, v.

We employ several ideas used by Raudenbush in the second of his above mentioned papers.

1. **The assumption.** Consider a form¶ G every term of which is divisible by some $u_i v_j$.** Let Σ be the set of all such forms G. Then Σ is a differential ideal, and is perfect. For, if a form has a term free of, say, every u_i, then every power of the form will have such a term.

---

* Presented to the Society, February 25, 1939.

† For the definition of differential field, and other terms, see H. W. Raudenbush, *Ideal theory and differential equations*, Transactions of this Society, vol. 36 (1934), pp. 361–368.

‡ Throughout the rest of this paper we shall use, as is customary, the term *form* for *differential polynomial*.

§ For differential fields of meromorphic functions this was essentially shown by J. F. Ritt in his book *Differential Equations from the Algebraic Standpoint*, American Mathematical Society Colloquium Publications, vol. 14, New York, 1932. See especially §§ 7, 77. Following the work of Ritt, Raudenbush treated the case of the general differential field of characteristic zero by purely algebraic methods. See Raudenbush, loc. cit.

‖ *On the analog for differential equations of the Hilbert-Netto theorem*, this Bulletin, vol. 42 (1936), pp. 371–373.

¶ For $\mathcal{F}$ we can use any differential field of characteristic zero.

** Subscripts denote derivatives.

It is easy to see that Σ is the perfect differential ideal generated by uv.

If Σ has a strong basis, then it has one consisting purely of forms u_iv_j.

Let

$$u_iv_j, \qquad i+j \leqq s, \tag{1}$$

be a strong basis for Σ, and let p be the associated positive integer. We work toward a contradiction.

We denote by α a positive integer to be fixed later.

Consider the set of all forms

$$u_{i_1}v_{j_1}\cdots u_{i_p}v_{j_p}, \qquad i_1+j_1+\cdots+i_p+j_p=\alpha. \tag{2}$$

Every such form has an expression $\sum_{\varrho=1}^{r}c_\varrho\left(\sum_{h=1}^{p}a_{\varrho h}u_{i_h}v_{j_h}\right)^p$, where r is some positive integer, and the c_ϱ and the $a_{\varrho h}$ are rational numbers.* Therefore, by our assumption on the nature of the basis (1) and the integer p, every form (2) is in the differential ideal generated by the forms (1).

Hence each form (2) is a linear combination, with coefficients in F, of forms

$$\begin{gathered}(u_iv_j)_ku_{i_1}v_{j_1}\cdots u_{i_{p-1}}v_{j_{p-1}},\\ i+j\leqq s,\ i+j+k+i_1+j_1+\cdots+i_{p-1}+j_{p-1}=\alpha.\end{gathered} \tag{3}$$

Since the forms (2) are all linearly independent over $\mathfrak{F}$, it follows that the number of distinct forms (2) cannot exceed the number of distinct forms (3).

We denote the number of distinct forms (2) by $R_{p,\alpha}$, and the number of distinct forms (3) by $Q_{p,\alpha}$. We thus have $R_{p,\alpha}\leqq Q_{p,\alpha}$.

In the next section we force the contradiction that $R_{p,\alpha}>Q_{p,\alpha}$ for α sufficiently large.

2. **The contradiction.** We consider those expressions (2) for which $i_1+j_1=\nu$, $(0\leqq\nu\leqq\alpha)$. The coefficient of $u_{i_1}v_{j_1}$ in (2) is then

$$u_{i_2}v_{j_2}\cdots u_{i_p}v_{j_p}, \qquad i_2+j_2+\cdots+i_p+j_p=\alpha-\nu. \tag{4}$$

The number of distinct forms (4) is $R_{p-1,\alpha-\nu}$, and therefore the number of distinct symbols† (4) is not less than $R_{p-1,\alpha-\nu}$. Since the number of expressions $u_{i_1}v_{j_1}$ with $i_1+j_1=\nu$ is $\nu+1$, the total number of sym-

* We can solve the equations $(w_1+\lambda w_2)^t=\sum_{i=0}^{t}C_{t,i}\lambda^iw_1^{t-i}w_2^i$, $(\lambda=1,\cdots,t+1)$, for $w_1w_2^{t-1}$, obtaining $w_1w_2^{t-1}=\sum_{\lambda=1}^{t+1}d_\lambda(w_1+\lambda w_2)^t$. Using this special case, we can show by induction that $w_1\cdots w_p=\sum_{\varrho=1}^{r}c_\varrho\left(\sum_{h=1}^{p}a_{\varrho h}w_h\right)^p$. Setting $w_h=u_{i_h}v_{j_h}$, we obtain the desired representation of the forms (2).

† Two distinct symbols (4) may represent the same form.

bols (2) is not less than $\sum_{\nu=0}^{\alpha}(\nu+1)R_{p-1,\alpha-\nu}$. But not more than $(p!)^2$ distinct symbols (2) can represent the same form. Hence

$$R_{p,\alpha} \geqq (p!)^{-2}\sum_{\nu=0}^{\alpha}(\nu+1)R_{p-1,\alpha-\nu}. \tag{5}$$

We now show that there exist positive numbers b_p, $(p=1, 2, \cdots)$, independent of α, such that

$$R_{p,\alpha} \geqq b_p(\alpha+1)^{2p-1}. \tag{6}$$

Obviously $R_{1,\alpha}=\alpha+1$, so that (6) holds for $p=1$. Suppose (6) holds for $p=m-1$. Then, by (5), using $[x]$ to denote the greatest integer not exceeding x, we have

$$\begin{aligned}
R_{m,\alpha} &\geqq (m!)^{-2}\sum_{\nu=0}^{\alpha}(\nu+1)b_{m-1}(\alpha-\nu+1)^{2m-3} \\
&\geqq (m!)^{-2}b_{m-1}\sum_{\nu=[\alpha/4]}^{[3\alpha/4]}(\nu+1)(\alpha-\nu+1)^{2m-3} \\
&\geqq (m!)^{-2}b_{m-1}\sum_{\nu=[\alpha/4]}^{[3\alpha/4]}([\alpha/4]+1)(\alpha/4+1)^{2m-3} \\
&\geqq (m!)^{-2}b_{m-1}(2[\alpha/4]+1)([\alpha/4]+1)(\alpha/4+1)^{2m-3} \\
&\geqq b_m(\alpha+1)^{2m-1},
\end{aligned}$$

where $b_m=(m!)^{-2}4^{-2m}b_{m-1}$. Thus (6) holds for all p.

We now consider those expressions (3) for which $i+j+k=\mu$, $(0\leqq\mu\leqq\alpha)$. The number of distinct expressions $(u_iv_j)_k$ with $i+j+k=\mu$ and with $i+j\leqq s$ does not exceed $(s+1)^2$. The coefficient of $(u_iv_j)_k$ in (3) is $u_{i_1}v_{j_1}\cdots u_{i_{p-1}}v_{j_{p-1}}$, $(i_1+j_1+\cdots+i_{p-1}+j_{p-1}=\alpha-\mu)$. Since the number of distinct forms of this kind is $R_{p-1,\alpha-\mu}$, we have for the total number of distinct forms (3):

$$Q_{p,\alpha} \leqq \sum_{\mu=0}^{\alpha}(s+1)^2R_{p-1,\alpha-\mu}. \tag{7}$$

We shall show that, for $p=1, 2, \cdots$,

$$R_{p,\alpha} \leqq (\alpha+1)^{2p-1}. \tag{8}$$

For since $R_{1,\alpha}=\alpha+1$, (8) holds for $p=1$. Suppose (8) holds for $p=m-1$. Looking at (2), it is easy to see that

$$R_{m,\alpha} \leqq \sum_{\nu=0}^{\alpha}(\nu+1)R_{m-1,\alpha-\nu}.$$

Therefore

$$R_{m,\alpha} \leqq \sum_{\nu=0}^{\alpha} (\nu+1)(\alpha-\nu+1)^{2m-3}$$

$$\leqq \sum_{\nu=0}^{\alpha} (\alpha+1)(\alpha+1)^{2m-3} = (\alpha+1)^{2m-1}.$$

Thus (8) holds for all p.

Using (8) in (7), we find

$$Q_{p,\alpha} \leqq (s+1)^2 \sum_{\mu=0}^{\alpha} (\alpha-\mu+1)^{2p-3},$$

so that

$$Q_{p,\alpha} \leqq (s+1)^2(\alpha+1)^{2p-2}.$$

Comparing this with (6), we see that, for α sufficiently large, $R_{p,\alpha} > Q_{p,\alpha}$.

COLUMBIA UNIVERSITY

RELATIONS AMONG THE FUNDAMENTAL SOLUTIONS OF THE GENERALIZED HYPERGEOMETRIC EQUATION WHEN $p=q+1$

II. LOGARITHMIC CASES*

F. C. SMITH

1. **Introduction.** In a previous paper [1], the author gave the relations among the non-logarithmic solutions of the equation

$$\left\{\prod_{t=1}^{q+1}(\theta+a_t)-\frac{1}{z}\prod_{t=1}^{q+1}(\theta+c_t-1)\right\}y=0 \tag{1}$$

where $\theta=z(d/dz)$ and where the a_t and c_t are any constants, real or complex, the only restriction being that one of the c_t must be equal to unity. If no two of the a_t or c_t are equal or differ by an integer, then equation (1) has $q+1$ linearly independent solutions about the point $z=0$ which may be written

$$Y_{0j}=z^{1-c_j}\prod_{t=1}^{q+1}\frac{\Gamma(1+c_t-c_j)}{\Gamma(1+a_t-c_j)}\sum_{n=0}^{\infty}\prod_{t=1}^{q+1}\frac{\Gamma(1+a_t-c_j+n)}{\Gamma(1+c_t-c_j+n)}z^n, \tag{2}$$
$$j=1,2,\cdots,q+1;\ |z|<1,$$

and $q+1$ linearly independent solutions about the point $z=\infty$ which may be written

$$Y_{\infty j}=z^{-a_j}\prod_{t=1}^{q+1}\frac{\Gamma(1-a_t+a_j)}{\Gamma(1-c_t+a_j)}\sum_{n=0}^{\infty}\prod_{t=1}^{q+1}\frac{\Gamma(1-c_t+a_j+n)}{\Gamma(1-a_t+a_j+n)}\frac{1}{z^n}, \tag{3}$$
$$j=1,2,\cdots,q+1;\ |z|>1.$$

If, however, we assume that

$$c_2-c_1=l_1;\ c_3-c_2=l_2;\ \cdots;\ c_r-c_{r-1}=l_{r-1}$$

where each l_v is zero or a positive integer and assume at the same time that none of these r c_t is equal to or differs from any of the a_t by an integer, then the author has shown that the first r of the solutions (2) are replaced by the following forms [2]:

$$Y_{0j}=\sum_{v=1}^{j}z^{1-c_v}\frac{(j-1)!}{(j-v)!}\left[\frac{\partial^{j-v}}{\partial w^{j-v}}z^wG_v(w,z)\right]_{w=0}, \tag{4}$$
$$j=1,2,\cdots,r;\ |z|<1,$$

where

* Presented to the Society, December 28, 1938.

$$G_v(w,z) = (-1)^{1-v-\sum_{t=1}^{v-1} t l_t}\left(\frac{\pi w}{\sin \pi w}\right)^{1-v} \prod_{t=1}^{q+1} \frac{\Gamma(1+c_t-c_1+w)}{\Gamma(1+a_t-c_1+w)}$$

$$\cdot \sum_{n=0}^{l_{v-1}-1} \prod_{t=1}^{v-1} \Gamma(c_v - c_t - w - n)\Gamma(1+a_t-c_v+w+n) \tag{5}$$

$$\cdot \prod_{t=v}^{q+1} \frac{\Gamma(1+a_t-c_v+w+n)}{\Gamma(1+c_t-c_v+w+n)}\left[(-1)^{v-1}z\right]^n$$

in which the -1 factor and the first product of the summation are missing when $v=1$, and in which we make the special definition $l_0=\infty$; moreover, we make the special convention that $G_v(w,z)=0$ if $l_{v-1}=0$. Similarly, if we assume that $a_1-a_2=k_1; a_2-a_3=k_2; \cdots; a_{s-1}-a_s=k_{s-1}$ where each k_v is zero or a positive integer and assume at the same time that none of these s a_t is equal to or differs from any of the c_t by an integer, then the first s of the solutions (3) are replaced by the following forms [2]:

$$Y_{\infty j} = \sum_{v=1}^{j} z^{-a_v} \frac{(j-1)!}{(j-v)!}\left[\frac{\partial^{j-v}}{\partial w^{j-v}} z^{-w} F_v(w,z)\right]_{w=0}, \tag{6}$$

$$j = 1, 2, \cdots, s; \ |z| > 1,$$

where

$$F_v(w,z) = (-1)^{1-v-\sum_{t=1}^{v-1} t k_t}\left(\frac{\pi w}{\sin \pi w}\right)^{1-v} \prod_{t=1}^{q+1} \frac{\Gamma(1-a_t+a_1+w)}{\Gamma(1-c_t+a_1+w)}$$

$$\cdot \sum_{n=0}^{k_{v-1}-1} \prod_{t=1}^{v-1} \Gamma(a_t - a_v - w - n)\Gamma(1-c_t+a_v+w+n) \tag{7}$$

$$\cdot \prod_{t=v}^{q+1} \frac{\Gamma(1-c_t+a_v+w+n)}{\Gamma(1-a_t+a_v+w+n)}\left[(-1)^{v-1}\frac{1}{z}\right]^n$$

in which we make special conventions of the same type as those made in connection with (5).

It is the purpose of this paper to develop the relations among the solutions of (1) when one or both of the two sets of solutions contain logarithmic members. The results of this paper generalize those of Mehlenbacher [3] and Lindelöf [4] who treated the case in which $q=1$.

2. **All Y_{0j} non-logarithmic.** In this case we may state the following theorem:

THEOREM 1. *If all of the solutions Y_{0j} of equation* (1) *are non-logarithmic in character while the first s of the solutions $Y_{\infty j}$ are logarith-*

mic, then the Y_{0j}, when extended analytically outside their circle of convergence, may be expressed linearly in terms of the $Y_{\infty j}$ in the following forms, it being understood throughout that $0<\arg z<2\pi$:

$$
\begin{aligned}
Y_{0j} = {} & \frac{e(a_s - c_j, 1)}{(s-1)!} \prod_{t=1}^{q+1} \frac{\Gamma(1 + c_t + c_j)}{\Gamma(1 + a_t - c_j)} \sum_{k=1}^{s} C_{s-1,k-1} P_j^{(s-k)}(0, 1) Y_{\infty k} \\
(8) \qquad & + \sum_{k=s+1}^{q+1} \left\{ e(a_k - c_j + 1, 1) \frac{\Gamma(c_j - a_k)\Gamma(1 + c_k - c_j)}{\Gamma(c_k - a_k)} \right. \\
& \left. \cdot \prod_{t=1, t\neq k}^{q+1} \frac{\Gamma(a_t - a_k)\Gamma(1 + c_t - c_j)}{\Gamma(c_t - a_k)\Gamma(1 + a_t - c_j)} Y_{\infty k} \right\}, \\
& \qquad\qquad\qquad\qquad\qquad\qquad j = 1, 2, \cdots, q+1,
\end{aligned}
$$

where $e(a_s - c_j, 1)$ and $e(a_k - c_j + 1, 1)$ are to be obtained from the definition

$$e(w, m) = \exp\{(1/2)i\pi[1 - (-1)^m]w\}$$

and where $P_j^{(s-k)}(0, 1)$ denotes the $(s-k)$th derivative with respect to w of the function

$$
\begin{aligned}
P_j(w, m) = {} & (-1)^{s + \sum_{t=1}^{s-1}(s+t+m)k_t} \cdot e(w, m) \left(\frac{\pi w}{\sin \pi w} \right)^s \\
(9) \qquad & \cdot \prod_{t=1}^{s} \frac{\Gamma^m(c_j - a_1 - w)\Gamma^m(w - c_j + a_1 + 1)}{\Gamma(c_t - a_1 - w)\Gamma(w - a_t + a_1 + 1)} \\
& \cdot \prod_{t=1}^{q+1} \frac{\Gamma(a_t - a_1 - w)}{\Gamma(c_t - a_1 - w)}
\end{aligned}
$$

*evaluated for $w=0$ and $m=1$.**

PROOF. The proof follows the same outline as that employed in the non-logarithmic cases [1]. Under present assumptions, the function

$$g(w) = \prod_{t=1}^{q+1} \frac{\Gamma(1 + a_t - c_j + w)}{\Gamma(1 + c_t - c_j + w)} \tag{10}$$

continues to have simple poles at the points

$$w = c_j - a_k - n - 1, \quad n = 0, 1, 2, \cdots; \ k = s+1, \cdots, q+1, \tag{11}$$

but now has poles of order $s-v+1$ at the points

* Throughout this paper, an expression of the form $P^{(k)}(w)$ indicates the kth derivative of $P(w)$ while an expression of the form $P^k(w)$ indicates the kth power of $P(w)$.

(12) $w = c_j - a_v - n - 1,\ n = 0, 1, 2, \cdots, k_{v-1} - 1; v = 1, 2, \cdots, s.$

Therefore, in the result previously obtained,* we must replace the terms for $k = 1, 2, \cdots, s$ by

$$(13) \qquad - z^{1-c_j} \prod_{t=1}^{q+1} \frac{\Gamma(1 + c_t - c_j)}{\Gamma(1 + a_t - c_j)} \sum_{v=1}^{s} \sum_{n=0}^{k_{v-1}-1} R_{n,v}$$

where $R_{n,v}$ denotes the residue of the function

$$(14) \qquad \frac{\pi(-z)^w g(w)}{\sin \pi w} = (-z)^w \Gamma(w)\Gamma(1-w) \prod_{t=1}^{q+1} \frac{\Gamma(1 + a_t - c_j + w)}{\Gamma(1 + c_t - c_j + w)}$$

at the point $w = c_j - a_v - n - 1$.

In order to compute $R_{n,v}$, consider

$$(15) \qquad R_{n,v} = \frac{1}{2\pi i} \int_{C_n} (-z)^w \Gamma(w)\Gamma(1-w) \prod_{t=1}^{q+1} \frac{\Gamma(1 + a_t - c_j + w)}{\Gamma(1 + c_t - c_j + w)}\, dw$$

where C_n^* surrounds the pole $w = c_j - a_v - n - 1$ of (14) but no other pole of (14). If in (15), we replace w by $-(w - c_j + a_v + n + 1)$, we get

$$(16) \qquad \begin{aligned} R_{n,v} = &- \frac{(-z)^{c_j - a_v - n - 1}}{2\pi i} \int_{C_0} (-z)^{-w} \Gamma(c_j - a_v - w - n - 1) \\ &\cdot \Gamma(w - c_j + a_v + n + 1) \prod_{t=1}^{q+1} \frac{\Gamma(a_t - a_v - w - n)}{\Gamma(c_t - a_v - w - n)}\, dw \end{aligned}$$

where C_0 surrounds the origin. By several applications of the relation

$$(17) \qquad \Gamma(1 - w) = \frac{\pi}{\Gamma(w) \sin \pi w}$$

we may change (16) to the form

$$(18) \qquad \begin{aligned} R_{n,v} = &- \frac{z^{c_j - a_v - 1} e(a_s - c_j, 1)(-1)^{1 - v - \sum_{t=1}^{v-1} k_t}}{[(-1)^{v-1} z]^n 2\pi i} \\ &\cdot \int_{C_0} \frac{z^{-w} P_j(w, 1)}{w^{s-v+1}} \left(\frac{\pi w}{\sin \pi w}\right)^{1-v} \prod_{t=1}^{q+1} \frac{\Gamma(1 - a_t + a_1 + w)}{\Gamma(1 - c_t + a_1 + w)} \\ &\cdot \prod_{t=1}^{v-1} \Gamma(a_t - a_v - w - n)\Gamma(1 - c_t + a_v + w + n) \\ &\cdot \prod_{t=v}^{q+1} \frac{\Gamma(1 - c_t + a_v + w + n)}{\Gamma(1 - a_t + a_v + w + n)}\, dw. \end{aligned}$$

* See [1], equation (5).

Therefore,

$$(19)\qquad \sum_{n=0}^{k_{\sigma-1}-1} R_{n,v} = -z^{c_j-a_\sigma-1}e(a_\sigma - c_j, 1)\int_{C_\sigma} \frac{z^{-w}P_j(w,1)F_v(w,z)}{w^{s-v+1}}\,dw.$$

When (19) is evaluated by the theorems of the calculus of residues [5], we get

$$(20)\qquad \sum_{n=0}^{k_{\sigma-1}-1} R_{n,v} = -\frac{z^{c_j-a_\sigma-1}e(a_\sigma - c_j, 1)}{(s-v)!}\left[\frac{\partial^{s-v}}{\partial w^{s-v}} z^{-w}P_j(w,1)F_v(w,z)\right]_{w=0}.$$

Therefore

$$\begin{aligned}(21)\qquad \sum_{v=1}^{s}\sum_{n=0}^{k_{\sigma-1}-1} R_{n,v} &= -z^{c_j-1}e(a_\sigma - c_j, 1)\sum_{v=1}^{s}\frac{z^{-a_\sigma}}{(s-v)!}\\ &\quad\cdot\sum_{h=0}^{s-v} C_{s-v,h}P_j^{(h)}(0,1)\left[\frac{\partial^{s-v-h}}{\partial w^{s-v-h}} z^{-w}F_v(w,z)\right]_{w=0}\\ &= -\frac{z^{c_j-1}e(a_\sigma - c_j, 1)}{(s-1)!}\sum_{k=1}^{s} C_{s-1,k-1}P_j^{(s-k)}(0,1)Y_{\sigma k}.\end{aligned}$$

Substituting (21) into (13), we get the final form of the replacements to be made in the result previously obtained. When these replacements are made, we obtain the desired result (8).

In a previous paper [1], the author gave the expression for the non-logarithmic $Y_{\infty j}$ in terms of the Y_{0j}. It remains, then, to develop the expression for the logarithmic $Y_{\infty j}$ in terms of the Y_{0j}. In this connection, we may state the following theorem:

THEOREM 2. *Under the conditions stated in Theorem* 1, *the logarithmic* $Y_{\infty j}$, *when extended analytically outside their circle of convergence, may be expressed linearly in terms of the* Y_{0j} *in the following forms, it being understood throughout that* $0<\arg(1/z)<2\pi$:

$$\begin{aligned}(22)\qquad Y_{\infty j} = -\sum_{k=1}^{q+1}\sum_{m=1}^{j}\Big\{&(m-1)!K_{mj}(L)e(c_k - a_m, m)\\ &\frac{\Gamma^m(c_k - a_m)\Gamma^m(1-c_k+a_m)}{\Gamma(c_k-a_k)}\prod_{t=1,t\neq k}^{q+1}\frac{\Gamma(c_k-c_t)}{\Gamma(c_k-a_t)}Y_{0k}\Big\},\\ &j = 1, 2, \cdots, s,\end{aligned}$$

where $K_{mj}(L)$ *is the quotient of the cofactor of the element in the mth row and jth column of the determinant*

$$(23)\quad \Delta_j(L) = \begin{vmatrix} L_1(0) & 0 & 0 & \cdots & 0 \\ L_2'(0) & L_2(0) & 0 & \cdots & 0 \\ L_3''(0) & 2L_3'(0) & L_3(0) & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \ddots \\ L_j^{(j-1)}(0) & C_{j-1,1}L_j^{(j-2)}(0) & C_{j-1,2}L_j^{(j-3)}(0) & \cdots & L_j(0) \end{vmatrix}$$

and $\Delta_j(L)$ itself, the $L_j(w)$ being defined by the formula

$$(24)\qquad L_j(w) = e(w,j)\left(\frac{\pi w}{\sin \pi w}\right)^j \prod_{t=1}^{q+1} \frac{\Gamma(1-c_t+a_1+w)}{\Gamma(1-a_t+a_1+w)}$$

and the cofactor of $L_1(0)$ in $\Delta_1(L)$ being defined equal to unity.

PROOF. In order to prove this theorem, we follow the same procedure as that used in the proof of a theorem due to Ford [6]. If in Ford's proof, we replace the integrand which he uses by

$$(25)\qquad \frac{(\pm z)^{-w}\pi^j}{\sin^j \pi w} \prod_{t=1}^{q+1} \frac{\Gamma(1-c_t+a_j+w)}{\Gamma(1-a_t+a_j+w)}, \quad j = 1, 2, \cdots, s,$$

where the upper or lower sign is to be taken according as j is even or odd, we obtain in place of Ford's final result

$$(26)\quad \sum_{n=-\infty}^{\infty} \frac{1}{2\pi i}\int_{C_n} \frac{(\pm z)^{-w}\pi^j}{\sin^j \pi w} \prod_{t=1}^{q+1} \frac{\Gamma(1-c_t+a_j+w)}{\Gamma(1-a_t+a_j+w)}\,dw \sim -\sum_{k=1}^{q+1}\sum_{n=0}^{\infty} S_{n,k,j}, \quad j = 1, \cdots, s;\ 0 < \arg(1/z) < 2\pi,$$

where C_n surrounds the pole $w=n$ of (25) but no other pole of (25), and where $S_{n,k,j}$ denotes the residue of (25) at the point $w=c_k-a_j-n-1$ which, under the present assumptions, is a simple pole of (25). If we replace w by $w+n$ on the left in (26), we obtain

$$(27)\quad \begin{aligned} &\sum_{n=-\infty}^{\infty} \frac{z^{-n}}{2\pi i}\int_{C_0} \frac{(\pm z)^{-w}\pi^j}{\sin^j \pi w} \prod_{t=1}^{q+1} \frac{\Gamma(1-c_t+a_j+w+n)}{\Gamma(1-a_t+a_j+w+n)}\,dw \\ &= \sum_{n=1}^{\infty} \frac{z^n}{2\pi i}\int_{C_0} \frac{(\pm z)^{-w}\pi^j}{\sin^j \pi w} \prod_{t=1}^{q+1} \frac{\Gamma(1-c_t+a_j+w-n)}{\Gamma(1-a_t+a_j+w-n)}\,dw \\ &\quad + \sum_{v=1}^{j} z^{a_j-a_v} \sum_{n=0}^{k_{v-1}-1} \frac{z^{-n}}{2\pi i}\int_{C_0} \frac{(\pm z)^{-w}\pi^j}{\sin^j \pi w} \\ &\quad \cdot \prod_{t=1}^{q+1} \frac{\Gamma(1-c_t+a_v+w+n)}{\Gamma(1-a_t+a_v+w+n)}\,dw \end{aligned}$$

where C_0 surrounds the origin. If we now apply (17) a number of times, the right side of (27) becomes

$$
\begin{aligned}
&\sum_{n=1}^{\infty} \frac{z^n(-1)^{jn+\sum_{i=1}^{j}(a_i-a_j-1)}}{2\pi i}\int_{C_0}(\pm z)^{-w} \\
&\qquad \cdot \prod_{i=1}^{j}\Gamma(a_i-a_j-w+n)\Gamma(1-c_i+a_j+w-n) \\
&\qquad \cdot \prod_{i=j+1}^{q+1}\frac{\Gamma(1-c_i+a_j+w-n)}{\Gamma(1-a_i+a_j+w-n)}\,dw \\
&\quad + \sum_{v=1}^{j}\frac{z^{a_j-a_v}}{2\pi i}\int_{C_0}\frac{z^{-w}L_j(w)F_v(w,z)}{w^{j-v+1}}\,dw.
\end{aligned}
\tag{28}
$$

But since the integrand of the first integral of (28) is analytic for $w=0$, the first summation vanishes. When the second term of (28) is evaluated by the theorems of the calculus of residues [5], we obtain

$$
\begin{aligned}
z^{a_j}\sum_{v=1}^{j}\frac{z^{-a_v}}{(j-v)!}\left[\frac{\partial^{j-v}}{\partial w^{j-v}}z^{-w}L_j(w)F_v(w,z)\right]_{w=0} \\
= \frac{z^{a_j}}{(j-1)!}\sum_{m=1}^{j}C_{j-1,m-1}L_j^{(j-m)}(0)Y_{\infty m}.
\end{aligned}
\tag{29}
$$

As to the right member of (26), since $w=c_k-a_j-n-1$ is a simple pole of (25), we may compute the residue $S_{n,k,j}$ without difficulty. When this residue and (29) are used in (26), we obtain

$$
\begin{aligned}
&\sum_{m=1}^{j}C_{j-1,m-1}L_j^{(j-m)}(0)Y_{\infty m} \\
&\sim -(j-1)!\sum_{k=1}^{q+1}\left\{e(c_k-a_j,j)\frac{\Gamma^j(c_k-a_j)\Gamma^j(1-c_k+a_j)}{\Gamma(c_k-a_k)}\right. \\
&\qquad \left.\cdot\prod_{i=1,i\neq k}^{q+1}\frac{\Gamma(c_k-c_i)}{\Gamma(c_k-a_i)}Y_{0k}\right\}, \qquad j=1,2,\cdots,s.
\end{aligned}
\tag{30}
$$

Since all $Y_{\infty m}$ and all Y_{0k} which appear in (30) are defined by series which converge for $|z|>1$, and $|z|<1$, respectively, we may replace $\sim$ by $=$ in (30). Although this gives us s equations, only the first j of these need be used in solving for a particular $Y_{\infty j}$. Solving these by Cramer's rule, we obtain the desired result (22).

3. **All $Y_{\infty j}$ non-logarithmic.** By means of proofs similar to those given in §2, we may establish the following theorems:

THEOREM 3. *If all of the solutions $Y_{\infty j}$ of equation* (1) *are non-logarithmic in character while the first r of the solutions Y_{0j} are logarithmic, then the logarithmic Y_{0j}, when extended analytically outside their circle of convergence, may be expressed linearly in terms of the $Y_{\infty j}$ in the following forms, it being understood throughout that* $0<\arg z<2\pi$:

$$Y_{0j} = -\sum_{k=1}^{q+1}\sum_{m=1}^{j}\Bigg\{(m-1)!K_{mj}(N)e(a_k-c_m,m) \cdot\frac{\Gamma^m(c_m-a_k)\Gamma^m(1-c_m+a_k)}{\Gamma(c_k-a_k)}\prod_{t=1,t\neq k}^{q+1}\frac{\Gamma(a_t-a_k)}{\Gamma(c_t-a_k)}Y_{\infty k}\Bigg\}, \tag{31}$$

$$j=1,2,\cdots,r,$$

where $K_{mj}(N)$ is defined as in Theorem 2 with $L_j(w)$ replaced by

$$N_j(w) = e(-w,j)\left(\frac{\pi w}{\sin \pi w}\right)^j\prod_{t=1}^{q+1}\frac{\Gamma(1+a_t-c_1+w)}{\Gamma(1+c_t-c_1+w)}. \tag{32}$$

THEOREM 4. *Under the conditions stated in Theorem 3, the $Y_{\infty j}$, when extended analytically outside their circle of convergence, may be expressed linearly in terms of the Y_{0j} in the following forms, it being understood throughout that* $0<\arg(1/z)<2\pi$:

$$\begin{aligned}Y_{\infty j} = {}& \frac{e(c_r-a_j,1)}{(r-1)!}\prod_{t=1}^{q+1}\frac{\Gamma(1-a_t+a_j)}{\Gamma(1-c_t+a_j)}\sum_{k=1}^{r}C_{r-1,k-1}Q_j^{(r-k)}(0,1)Y_{0k}\\ &+\sum_{k=r+1}^{q+1}\Bigg\{e(c_k-a_j-1,1)\frac{\Gamma(c_k-a_j)\Gamma(1-a_k+a_j)}{\Gamma(c_k-a_k)}\\ &\cdot\prod_{t=1,t\neq k}^{q+1}\frac{\Gamma(c_k-c_t)\Gamma(1-a_t+a_j)}{\Gamma(c_k-a_t)\Gamma(1-c_t+a_j)}Y_{0k}\Bigg\},\end{aligned} \tag{33}$$

$$j=1,2,\cdots,q+1,$$

where $Q_j^{(r-k)}(0,1)$ denotes the $(r-k)$th derivative with respect to w of the following function, evaluated for $w=0$ and $m=1$:

$$Q_j(w,m) = (-1)^{r+\sum_{t=1}^{r-1}(r+t+m)l_t}e(-w,m)\left(\frac{\pi w}{\sin \pi w}\right)^r \cdot\prod_{t=1}^{r}\frac{\Gamma^m(c_1-a_j-w)\Gamma^m(w-c_1+a_j+1)}{\Gamma(c_1-a_t-w)\Gamma(w-c_1+c_t+1)}\prod_{t=r+1}^{q+1}\frac{\Gamma(c_1-c_t-w)}{\Gamma(c_1-a_t-w)}. \tag{34}$$

4. **Logarithmic members in both Y_{0j} and $Y_{\infty j}$.** The procedure used in the proof of Theorem 2 may be used again in this case. The only additional difficulties arise from the fact that the residue terms like $S_{n,k,j}$ of (26) do not all come from simple poles. But the steps in the

proofs of the following theorems are the same as those in the proof of Theorem 2.

THEOREM 5. *If the first r of the solutions Y_{0j} and the first s of the solutions $Y_{\infty j}$ of equation* (1) *are logarithmic in character, then the logarithmic Y_{0j}, when extended analytically outside their circle of convergence, may be expressed linearly in terms of the $Y_{\infty j}$ in the following forms, it being understood throughout that* $0<\arg z<2\pi$:

$$Y_{0j} = \sum_{m=1}^{j} (m-1)!K_{mj}(N)\left\{\frac{e(a_s - c_m, m)}{(s-1)!}\sum_{k=1}^{s} C_{s-1,k-1}P_m^{(s-k)}(0,m)Y_{\infty k}\right.$$

$$(35)\qquad -\sum_{k=s+1}^{q+1} e(a_k - c_m, m)\frac{\Gamma^m(c_m - a_k)\Gamma^m(1 - c_m + a_k)}{\Gamma(c_k - a_k)}$$

$$\left.\cdot\prod_{t=1,t\neq k}^{q+1}\frac{\Gamma(a_t - a_k)}{\Gamma(c_t - a_k)}\,Y_{\infty k}\right\}, \qquad j = 1, 2, \cdots, r.$$

THEOREM 6. *Under the conditions stated in Theorem* 5, *the logarithmic $Y_{\infty j}$, when extended analytically outside their circle of convergence, may be expressed linearly in terms of the Y_{0j} in the following forms, it being understood throughout that* $0<\arg(1/z)<2\pi$:

$$Y_{\infty j} = \sum_{m=1}^{j} (m-1)!K_{mj}(L)\left\{\frac{e(c_r - a_m, m)}{(r-1)!}\sum_{k=1}^{r} C_{r-1,k-1}Q_m^{(r-k)}(0,m)Y_{0k}\right.$$

$$(36)\qquad -\sum_{k=r+1}^{q+1} e(c_k - a_m, m)\frac{\Gamma^m(c_k - a_m)\Gamma^m(1 - c_k + a_m)}{\Gamma(c_k - a_k)}$$

$$\left.\cdot\prod_{t=1,t\neq k}^{q+1}\frac{\Gamma(c_k - c_t)}{\Gamma(c_k - a_t)}\,Y_{0k}\right\}, \qquad j = 1, 2, \cdots, s.$$

REFERENCES

1. F. C. Smith, *Relations among the fundamental solutions of the generalized hypergeometric equation when $p=q+1$. I. Non-logarithmic cases*, this Bulletin, vol. 44 (1938), pp. 429–433.

2. F. C. Smith, *On the logarithmic solutions of the generalized hypergeometric equation when $p=q+1$*, this Bulletin, vol. 45 (1939), pp. 629–636.

3. L. E. Mehlenbacher, *Interrelations of the fundamental solutions of the hypergeometric equation; logarithmic case*, American Journal of Mathematics, vol. 60 (1938), pp. 120–128.

4. Ernst Lindelöf, *Sur l'intégration de l'équation de Kummer*, Acta Societatis Scientiarum Fennicae, (1), vol. 19 (1893), pp. 3–31.

5. W. B. Ford, *The Asymptotic Developments of Functions Defined by Maclaurin Series*, University of Michigan Press, 1936, pp. 5–10.

6. C. V. Newsom, *On the character of certain entire functions in distant portions of the plane*, American Journal of Mathematics, vol. 60 (1938), p. 562, equation (6).

THE COLLEGE OF ST. FRANCIS

THE ANALOGUE OF THE MOEBIUS GROUP OF CIRCULAR TRANSFORMATIONS IN THE KASNER PLANE*

JOHN DE CICCO

1. **Introduction.** We shall begin by giving some fundamental definitions. In this paper, by a curve of the plane π we shall mean a differential element of the third order in the plane π. A *simple horn-set* consists of all the curves (third order differential elements) in the plane π which possess a common point and a common direction. Let x denote the curvature and $y = dx/ds$ the rate of variation of the curvature per unit length of arc s of any curve of a simple horn-set at the common point. Then any curve of a simple horn-set is given by an ordered pair of numbers (x, y). From this, it follows that a simple horn-set of the plane π is a two-dimensional space, called the *Kasner plane* K_2, where any point of K_2 is a curve (x, y) of the simple horn-set. Thus to a given simple horn-set of the plane π, there is associated an auxiliary plane, called the Kasner plane K_2, such that any given point of the Kasner plane represents a curve of the simple horn-set whose curvature and rate of variation of the curvature per unit length of arc at the common point are the abscissa x and the ordinate y of the given point.

Kasner has shown that the group of conformal transformations in the plane π operating on the curves of a simple horn-set induces the *three-parameter group* G_3:

$$X = mx + h, \qquad Y = m^2y + k, \tag{1}$$

where $m \neq 0$, h, k are constants, from the points (x, y) to the points (X, Y) of the Kasner plane.† The Kasner plane is thus an affine plane.

A line consists of the ∞^1 points $C(x, y)$ of the Kasner plane which satisfy a linear equation in x and y with the coefficients of x and y not both zero. With respect to the group G_3, the lines of the Kasner plane may be classified into three distinct types.

(a) A *general line* is a line whose equation is of the form $y = px + r$, where $p \neq 0$ and r are constants.

(b) An *infinite line* is a line whose equation is of the special form $y = \text{const.}$

* Presented to the Society, February 25, 1939.

† Kasner, *Conformal geometry*, Proceedings of the Fifth International Congress of Mathematicians, Cambridge, 1912, vol. 2. Kasner and Comenetz, *Conformal geometry of horn angles*, Proceedings of the National Academy of Sciences, vol. 22 (1936), pp. 303–309.

(c) A *zero line* is a line whose equation is of the special form $x=\text{const.}$

Two lines possess a simple common point or are parallel. A line parallel to neither the x nor the y axis is a general line, whereas a line parallel to the x (or y) axis is an infinite (or zero) line.

The Kasner plane is a special type of Finsler two-space with the special Finsler metric $ds=dx^2/dy$. The finite form of this Finsler metric is given by (2) below. The infinite lines and the zero lines are the minimal lines of the Kasner plane. By (2), the distance between any two distinct points of an infinite (or zero) line is infinite (or zero). These statements give the reasons for the names—infinite lines and zero lines.

Under G_3, two points of the Kasner plane possess the unique invariant*

$$M_{12}=M(C_1, C_2)=\frac{(x_2-x_1)^2}{(y_2-y_1)}, \tag{2}$$

$$M_{21}=-M_{12}.$$

Under G_3, two general lines possess the invariant†

$$\alpha_{12}=\alpha(L_1, L_2)=p_2/p_1, \qquad \alpha_{21}=1/\alpha_{12}. \tag{3}$$

In the Kasner plane, M_{12} is the *distance or measure* between two points and α_{12} is the *dihorn angle* between two general lines.

The set of ∞^1 points $C(x, y)$ of the Kasner plane which satisfy an equation of the form

$$a_0x^2+a_1x+a_2y+a_3=0, \tag{4}$$

where at least one of the coefficients of the variables is not zero, is called a *parabolic-circle*. With respect to the group G_3, the parabolic-circles may be classified into four distinct types.

(a) *The proper parabolic-circles.* Any parabolic-circle whose equation is of the form (4) where $a_0a_2\neq 0$ is called a *proper parabolic-circle*. This may also be defined as the set of ∞^1 points $C(x, y)$ which are at a constant distance m from a fixed point $C_0(x_0, y_0)$. The distance m is called the *radius* and the fixed point C_0 is called the *center* of the parabolic-circle. The center and radius of the proper parabolic-circle (4) are

* See the preceding papers. M_{12} is the conformal measure of the horn angle formed by the two curves C_1 and C_2.

† Kasner, *Trihornometry: a new chapter of conformal geometry*, Proceedings of the National Academy of Sciences, vol. 23 (1937), pp. 337–341.

$$(5) \qquad x_0 = -a_1/2a_0, \qquad y_0 = -a_2/a_2 + a_1^2/4a_0a_2,$$
$$m = -a_2/a_0.$$

For a proper parabolic-circle, the center is a finite point and the radius is a finite nonzero number.

(b) *The general linear parabolic-circles.* Any parabolic-circle whose equation is of the form (4) where $a_0=0$, $a_1\neq 0$, $a_2\neq 0$ is called a general linear parabolic-circle. It is of course a general line. We may consider this as the limiting form of a proper parabolic-circle when the center goes to infinity and the radius becomes infinite.

(c) *The infinite linear parabolic-circles.* Any parabolic-circle whose equation is of the form (4) where $a_0=0$, $a_1=0$, $a_2\neq 0$, is called an infinite linear parabolic-circle. It is an infinite line. We may consider this as the limiting form of a proper parabolic-circle when the center remains a finite point and the radius becomes infinite.

(d) *The zero parabolic-circles.* Any parabolic-circle whose equation is of the form (4) where $a_2=0$ is called a zero parabolic-circle. It consists of none (imaginary), one, or two zero lines. This may be considered as the limiting form of a proper parabolic-circle when the center remains a finite point or goes to infinity and the radius becomes zero.

From (4), it is seen that *the differential equation of all parabolic-circles is* $y'''=0$.

The problem of this paper is to obtain the group of all point transformations with nonvanishing jacobian of the Kasner plane which convert every parabolic-circle into a parabolic-circle.* Of course the Kasner group G_3 as given by equations (1) is a subgroup of this new group. It is found that *this parabolic-circular group of point transformations is a continuous seven-parameter group* G_7. This group G_7 is the analogue of the Moebius group of circular transformations, which is a mixed six-parameter group generated by inversions. It can be shown that our group G_7 is isomorphic to the Laguerre group of circular line transformations. This is defined as the group of all line transformations which convert every circle into a circle.

We obtain the following fundamental result. *Let a point transformation T of the Kasner plane convert more than* $2\infty^2$ *parabolic-circles into parabolic-circles. That is, let T convert* $3\infty^2$ *parabolic-circles into parabolic-circles. Then T must be a parabolic-circular transformation.* This gives us a *minimal* characterization of the parabolic-circular

* Lie has shown that *the group of contact transformations which leave invariant the differential equation* $y'''=0$ *is a ten-parameter group* G_{10}. See Lie-Scheffers, *Berührungstransformationen*, p. 241. It may be proved that *if a contact transformation converts* $4\infty^2$ *vertical parabolas into vertical parabolas, then it must be a transformation of the Lie group* G_{10}. The same results are valid for circles.

group G_7 of the Kasner plane.* The analogous minimal characterization for the Laguerre group is the following: *Let a line transformation T convert more than $2\infty^2$ circles into circles. Then the line transformation T must be a circular line transformation of the Laguerre group.*

We close our paper with a characterization of the group of all point (or line) transformations in the plane π by means of the parabolic-circles of the Kasner plane.

2. **Parabolic-circles into parabolic-circles.** Let T be the point transformation with nonvanishing jacobian of the Kasner plane

$$X = X(x, y), \qquad Y = Y(x, y). \tag{6}$$

We wish to find out which parabolic-circles become parabolic-circles under T. In that event, the differential equation $y'''=0$ must be carried into $Y'''=0$ under the transformation T as given by (6). From (6) and this last fact, we see that the parabolic-circles which become parabolic-circles under T must satisfy the ordinary differential equation of the second order (of second degree in y'')

$$\lambda_x + p\lambda_y + p'(\lambda_p + J_x + pJ_y) - \frac{3(\lambda + p'J)(X_{xx} + 2pX_{xy} + p^2X_{yy} + p'X_y)}{X_x + pX_y} = 0, \tag{7}$$

where

$$\begin{aligned} \lambda &= (X_x + pX_y)(Y_{xx} + 2pY_{xy} + p^2Y_{yy}) \\ &\quad - (Y_x + pY_y)(X_{xx} + 2pX_{xy} + p^2X_{yy}), \\ J &= X_xY_y - X_yY_x, \end{aligned} \tag{8}$$

and $p=y'$. From (7) and (8), we obtain the following theorem.

THEOREM 1. *Let the point transformation T be such that $X_y \neq 0$. That is, let T not convert every zero line into a zero line. Then under T, there are at most either* (a) *two distinct two-parameter families* ($2\infty^2$) *of parabolic-circles, or* (b) *one two-parameter family* (∞^2) *of parabolic-circles (in this case, the two families coincide), or* (c) *no parabolic-circles (in this case, the two families are imaginary), which become parabolic-circles.*

* See the analogous characterizations for collineations and for the Moebius group, Kasner, *The characterization of collineations*, this Bulletin, vol. 9 (1903), pp. 545–546. A refined analysis of the continuity requirements of this problem is given in the Columbia dissertation of Prenowitz, Transactions of this Society, vol. 38 (1935), pp. 564–599. Also, Kasner and De Cicco, *Characterization of the Moebius group of circular transformations*, Proceedings of the National Academy of Sciences, vol. 25 (1939).

Let the point transformation T convert every zero line into a zero line. Then $X_y=0$ (hence $X_{xy}=X_{yy}=0$). We then find that (7) and (8) become respectively

$$\lambda_x + p\lambda_y + p'(\lambda_p + J_x + pJ_y) - 3(\lambda + p'J)X_{xx}/X_x = 0, \tag{9}$$

where

$$\lambda = X_x(Y_{xx} + 2pY_{xy} + p^2Y_{yy}) - X_{xx}(Y_x + pY_y), \quad J = X_xY_y. \tag{10}$$

From (9) and (10), we find that the coefficient of p' in (9) is

$$3(X_xY_{xy} - X_{xx}Y_y) + 3pX_xY_{yy}. \tag{11}$$

Before proceeding, we wish to introduce some definitions. Let two curves of the Kasner plane intersect in a given point and let their slopes at the given point be p_1 and p_2. The difference of the slopes p_2-p_1 is called the *slope-difference* between the two curves at the given point. Any point transformation of the Kasner plane which carries equal slope-differences at one point or at distinct points into equal slope-differences is termed a *magnilong transformation*. Obviously any such transformation is of the form

$$X = X(x), \qquad Y = \gamma X_x y + \beta(x), \tag{12}$$

where X and β are functions of x only and γ is a constant. We observe that (12) also converts every zero line into a zero line. The set of all magnilong transformations forms a group. It can be shown that this group of transformations is isomorphic to the group of line transformations, any one of which magnifies the distance between the two points of contact on the common tangent line of any two curves. This last statement gives the reason for the name magnilong.

From (9), (10), and (11), we obtain the following result.

THEOREM 2. *Let the point transformation T be such that $X=X(x)$ and $Y\neq\gamma X_xy+\beta(x)$, where X and β are functions of x only and γ is a constant. This means that T is not a magnilong transformation but T converts every zero line into a zero line. Then under T there is at most one two-parameter family (∞^2) of parabolic-circles which become parabolic-circles. Of course under T there are ∞^1 parabolic-circles which become parabolic-circles. These are the ∞^1 zero lines (which become zero lines).*

Next let T be the magnilong transformation (12). Then (9) and (10) become

$$\lambda_x + p\lambda_y - 3\lambda X_{xx}/X_x = 0, \tag{13}$$

where

$$\lambda = (X_x\beta_{xx} - X_{xx}\beta_x) + \gamma y(X_xX_{xxx} - X_{xx}^2) + p\gamma X_xX_{xx},$$
$$J = \gamma X_x^2. \tag{14}$$

Substituting (14) into (13), we obtain

$$\frac{\partial}{\partial x}(X_x\beta_{xx} - X_{xx}\beta_x) - \frac{3X_{xx}}{X_x}(X_x\beta_{xx} - X_{xx}\beta_x)$$
$$+ \gamma y\left[\frac{\partial}{\partial x}(X_xX_{xxx} - X_{xx}^2) - \frac{3X_{xx}}{X_x}(X_xX_{xxx} - X_{xx}^2)\right] \tag{15}$$
$$+ \gamma p[2X_xX_{xxx} - 3X_{xx}^2] = 0.$$

THEOREM 3. *Let the point transformation T be such that $X = X(x) \neq (ax+b)/(cx+d)$ and $Y = \gamma X_x y + \beta(x)$ where X and β are functions of x and γ, a, b, c, d are constants. This means that T is a magnilong transformation, but that it does not preserve the cross ratio*

$$(x_3 - x_1)(x_4 - x_2)/(x_3 - x_2)(x_4 - x_1)$$

of every four zero lines. There are at most two one-parameter families $(2\,\infty^1)$ of parabolic-circles which are converted into parabolic-circles. Of course, one of the two one-parameter families of parabolic-circles is the pencil of zero lines (which become zero lines).

Next let the point transformation T be of the form

$$X = \frac{ax + b}{cx + d}, \qquad Y = \frac{\gamma(ad - bc)}{(cx + d)^2}\, y + \beta(x), \tag{16}$$

where $\beta(x)$ is a function of x only and γ, a, b, c, d are constants. Upon substituting these values into (15), we find that the coefficient of y is zero and we have

$$\frac{\dfrac{\partial}{\partial x}(X_x\beta_{xx} - X_{xx}\beta_x)}{X_x\beta_{xx} - X_{xx}\beta_x} = \frac{3X_{xx}}{X_x}. \tag{17}$$

From (17), we obtain the following theorem.

THEOREM 4. *Let the point transformation T be such that*

$$X = \frac{ax + b}{cx + d}, \qquad Y = \frac{ey}{(cx + d)^2} + \beta(x), \tag{18}$$

where $\beta(x) \neq (fx^2+gx+h)/(cx+d)^2$ is a function of x only and a, b, c, d, e, f, g, h are constants. That is, T is a magnilong transformation such

that the cross ratio $(x_3-x_1)(x_4-x_2)/(x_3-x_2)(x_4-x_1)$ *between every four zero lines is preserved; but it does not convert every parabolic-circle into a parabolic-circle. Then the only parabolic-circles which under T are converted into parabolic-circles are the* ∞^1 *zero lines (which of course become zero lines).*

THEOREM 5. *The group of point transformations T of the Kasner plane which convert every parabolic-circle into a parabolic-circle is given by the equations*

$$X = \frac{ax+b}{cx+d}, \qquad Y = \frac{ey+fx^2+gx+h}{(cx+d)^2}, \tag{19}$$

where a, b, c, d, e, f, g, h are constants. Thus our group is a continuous seven-parameter group G_7.

From Theorems 1, 2, 3, 4, 5, we obtain the following conclusion.

THEOREM 6. *Let a point transformation T of the Kasner plane convert more than* $2\infty^2$ *parabolic-circles into parabolic-circles. That is, let T convert* $3\infty^2$ *parabolic-circles into parabolic-circles. Then T must be a parabolic-circular transformation.*

THEOREM 7. *The group* G_7 *of parabolic-circular point transformations has the following two properties:*

(a) *Every transformation of* G_7 *is a magnilong transformation. Hence every zero line is carried into a zero line.*

(b) *The cross ratio* $(x_3-x_1)(x_4-x_2)/(x_3-x_2)(x_4-x_1)$ *between every four zero lines is preserved.*

We note that if in (19) we take e to be the quantity $ad-bc$, the resulting six-parameter group may be written in the compact form

$$Z = \frac{\alpha z+\beta}{\gamma z+\delta}, \tag{20}$$

where each of the quantities $z, Z, \alpha, \beta, \gamma, \delta$ represents a *dual complex number* of the form $a+jb$ where $j^2=0$, and a and b are real numbers. Of course the transformation (20) is from the point $z=x+jy$ to the point $Z=X+jY$ of the Kasner plane.

By means of the above, we see that any transformation of the parabolic-circular group G_7 is the product of a linear fractional transformation of the form (20) followed by the transformation

$$X = x, \qquad Y = ey, \tag{21}$$

where e is an arbitrary real constant.

3. **The characterization of the group of all point (line) transformations in the plane π by the parabolic-circles of the Kasner plane.** It is found that the group of all contact transformations in the plane π induces an eight-parameter group G_8 of transformations in the Kasner plane. (See Kasner, *Characterization of the conformal group and the equilong group by horn angles*, Duke Mathematical Journal, vol. 4 (1938), pp. 95–106.) By examining this eight-parameter group G_8, we find the following two results.

THEOREM 8. *The subgroup G_6 of the parabolic-circular group G_7, which is given in the Kasner plane by the equations*

$$X = ax + b, \qquad Y = ey + fx^2 + gx + h, \tag{22}$$

is induced by the group of all point transformations in the plane π operating on the curves $C(x, y)$ (third order differential elements) of a simple horn-set.

THEOREM 9. *If a contact transformation of lineal elements in the plane π operating on the curves $C(x, y)$ (third order differential elements) of any simple horn-set converts every parabolic-circle of the Kasner plane into a parabolic-circle, then the contact transformation in the plane π must be a point transformation.*

Theorems 8 and 9 give us a new characteristic property of the group of all point transformations among all curve preserving transformations.

Let us now think of x as being the radius of curvature and $y = dx/du$ as being the rate of variation of the radius of curvature per unit radian measure of the angle u that the tangent line makes with a fixed line of the curve C of a simple horn-set at the common point. The Kasner group G_3 as given by the equations (1) will be induced by the group of equilong transformations. (Any transformation of this group is any line transformation such that the distance between the points of contact on the common tangent line of any two curves is preserved.) The distance M_{12} will be the measure of the horn angle in equilong geometry. (See Kasner, *Equilong invariants and convergence proofs*, this Bulletin, vol. 23 (1917), pp. 341–347.)

Theorems 8 and 9 will be true if we substitute the word "line" for the word "point." That is, these theorems give us a characteristic property of the group of all line transformations among all curve preserving transformations.

BROOKLYN COLLEGE

COMBINATORIC INTERPRETATION OF A FORMULA FOR THE nTH DERIVATIVE OF A FUNCTION OF A FUNCTION

I. OPATOWSKI

Let $f(x)=F(\phi(x))$. The formula of Faa' di Bruno* states that

$$\frac{d^n f}{dx^n}=\sum_{k=1}^{n}\frac{d^k F}{d\phi^k}\sum_{s}A_{ks}\prod_{r}^{(s)}\left(\frac{d^{k_r}\phi}{dx^{k_r}}\right)^{i_r}, \tag{1}$$

where $\prod_r^{(s)}$ is taken for *a system* (s) of positive integral solutions i_r, k_r of the equations $\sum_r i_r=k$, $\sum_r k_r i_r=n$ and $\sum_s$ is taken for *all such systems*. The factor A_{ks} is equal to $n!\left[\prod_r i_r!(k_r!)^{i_r}\right]^{-1}$. From a recent result of H. S. Wall† we have therefore *that the numerical factor* A_{ks} *in* (1) *is equal to the number of ways that n different objects can be placed in* $k=\sum_r i_r$ *compartments,* k_r *in each of* i_r *compartments, without regard to the order of arrangement of the compartments.*

H. S. Wall expressed the nth derivative of $f(x)$ in terms of logarithmic derivatives of $f(x)$.‡ Putting $F(\phi)=e^{\phi}$, $\phi(x)=\log f(x)$ in (1) his formula appears as a particular case of (1).

For *functions of many variables* $f(x)=F(\phi_1(x),\ \phi_2(x),\ \cdots,\ \phi_n(x))$ the formula of F. G. Teixeira§ states that

$$\frac{d^n f}{dx^n}=\sum_{k=1}^{n}\sum_{\sigma}\frac{\partial^k F}{\prod_t^{(\sigma)}\partial x_t^{a_t}}\sum_{s}A_{ks\sigma}\prod_{tr}^{(s)}\left(\frac{d^{k_{tr}}\phi_t}{dx_t^{k_{tr}}}\right)^{i_{tr}},$$

where $\prod_t^{(\sigma)}$ is taken for a system (σ) of nonnegative integral solutions a_t of the equation $\sum_t a_t=k$ and $\sum_\sigma$ is taken for all such systems; $\prod_{tr}^{(s)}$ is taken for a system (s) of positive integral solutions k_{tr}, i_{tr} of the equations $\sum_r i_{tr}=a_t$, $\sum_{tr} i_{tr}k_{tr}=n$ and $\sum_s$ is taken for all such systems. $A_{ks\sigma}$ is equal to $n!\left[\prod_{tr} i_{tr}!(k_{tr}!)^{i_{tr}}\right]^{-1}$. $A_{ks\sigma}$ has therefore the same combinatoric meaning as A_{ks}.

JOHNS HOPKINS UNIVERSITY

* F. Faa' di Bruno, *Sullo sviluppo delle funzioni*, Annali di Scienze Matematiche e Fisiche di Tortolini, vol. 6 (1855), pp. 479–480. For a bibliography on the subject see A. Voss, Encyclopädie der mathematischen Wissenschaften, II A 2, p. 88; E. Pascal, *Esercizi Critici di Calcolo*, Milano, 1921, 3d edition, pp. 111–112. See also L. S. Dederick, Annals of Mathematics, (2), vol. 27 (1926), pp. 385–394.

† H. S. Wall, *On the nth derivative of* $f(x)$, this Bulletin, vol. 44 (1938), pp. 395–397; see the theorem p. 395, and formula (13).

‡ Wall, loc. cit., formula (1).

§ F. G. Teixeira, *Sur les dérivées d'ordre quelconque*, Giornale di Matematica di Battaglini, vol. 18 (1880), p. 306.

MEANS OF ITERATED TRANSFORMATIONS IN REFLEXIVE VECTOR SPACES*

EDGAR R. LORCH

If U is a unitary transformation in a Hilbert space, the convergence of the sequence of transformations $\{(1/n)\sum_{s=0}^{n-1}U^s\}$ to a projection P has been established by von Neumann in his proof of the quasi-ergodic hypothesis.† Recently, the question of the existence of mean ergodic theorems in more general spaces than Hilbert space has received attention. One is here concerned with a bounded linear transformation V of bound not greater than 1, that is, $|V| \leq 1$, defined over a Banach space $\mathfrak{B}$, or more generally with a transformation V whose iterates V^n, $n=0, 1, 2, \cdots$, $(V^0=I)$ are uniformly bounded: $|V^n| \leq K$. The problem then is to characterize the spaces $\mathfrak{B}$ for which the arithmetic means $T_n=(1/n)\sum_{s=0}^{n-1}V^s$ converge, and to examine the nature of the limiting transformation. The case $\mathfrak{B}=L^p$, $p>1$, has been treated by F. Riesz.‡ Very recently, Garrett Birkhoff, considering transformations of bound not greater than 1, has extended Riesz' results to uniformly convex spaces.§ Our Theorem 2 includes all these results as special cases.‖ The methods we use resemble closely at certain points those used in a Hilbert space. The properties of reflexive spaces¶ to which appeal is made are simple and are examples of the elegant reciprocity relations which characterize those spaces (see, for instance, Theorem 1).

The adjoint space of a space $\mathfrak{B}$ is denoted by $(\mathfrak{B})$. The adjoint of

---

* Presented to the Society, September 8, 1939. Abstracted in this Bulletin, abstract 45-7-270 (1939).

† J. von Neumann, *Proof of the quasi-ergodic hypothesis*, Proceedings of the National Academy of Sciences, vol. 18 (1932), pp. 70–82.

‡ F. Riesz, *Some mean ergodic theorems*, Journal of the London Mathematical Society, vol. 13 (1938), pp. 274–278.

§ Garrett Birkhoff, this Bulletin, abstract 45-1-77 (1939); Duke Mathematical Journal, vol. 5 (1939), pp. 19–20; J. A. Clarkson, *Uniformly convex spaces*, Transactions of this Society, vol. 40 (1936), pp. 396–414.

‖ The results of this note were communicated orally by the author to Professor B. O. Koopman in the spring of 1938. It had been planned originally to delay announcement and publication until certain other researches had been completed. But since theorems of the type treated here have received considerable attention recently, it was decided to publish this proof separately without further delay.

¶ See the author's *On a calculus of operators in reflexive spaces*, Transactions of this Society, vol. 45 (1939), pp. 217–234. A discussion of the elementary properties of manifolds in reflexive spaces and of projections in any Banach space is found in Chapter II.

$(\mathfrak{B})$ is then $((\mathfrak{B}))$. The relation $\mathfrak{B} \subset ((\mathfrak{B}))$ expresses the fact that a certain obvious isometric mapping carries $\mathfrak{B}$ into a subset of $((\mathfrak{B}))$. If $\mathfrak{B} = ((\mathfrak{B}))$, the space $\mathfrak{B}$ is called *reflexive*. More precisely, if beginning with a space $\mathfrak{B}$, a $\Phi \in ((\mathfrak{B}))$ be chosen, then $\mathfrak{B}$ is reflexive if there exists an $f \in \mathfrak{B}$ such that $\Phi(F) = F(f)$ for all $F \in (\mathfrak{B})$. The spaces L^p, $p > 1$, are well known to be reflexive. The fact that any uniformly convex space is reflexive has recently been demonstrated.* If $\mathfrak{M}$ is a closed linear manifold (c.l.m.) in $\mathfrak{B}$, $\mathfrak{M}^\perp$ denotes the c.l.m. in $(\mathfrak{B})$ consisting of all $F \in (\mathfrak{B})$ orthogonal to every $f \in \mathfrak{M}$. If $\mathfrak{B}$ is reflexive, $\mathfrak{M}^{\perp\perp} = \mathfrak{M}$.†

A bounded linear transformation (b.l.t.) P in any $\mathfrak{B}$ is called a *projection* if $P^2 = P$. If P is a projection, there exist two c.l.m. $\mathfrak{M}$ and $\mathfrak{N}$ with $Pg = g$ if $g \in \mathfrak{M}$, $Ph = 0$ if $h \in \mathfrak{N}$ and for every $f \in \mathfrak{B}$, unique $g \in \mathfrak{M}$, $h \in \mathfrak{N}$ exist such that $f = g + h$. Conversely, manifold pairs of the type just described generate a projection.†

Let $\overline{T}$ be the adjoint of a b.l.t. T. Then it is known that $|\overline{T}| = |T|$. The principal result of this note will be proved with the assistance of the following theorem.

THEOREM 1. *Let $\mathfrak{B}$ be a reflexive vector space. Let T be a bounded linear transformation in $\mathfrak{B}$. Let $\mathfrak{M}$ denote the closed linear manifold of elements f for which $Tf = 0$; let $\mathfrak{N}$ denote the closed linear manifold spanned by the elements Tf, $f \in \mathfrak{B}$. Let $\overline{T}$ be the adjoint of T, and let $(\mathfrak{M})$ and $(\mathfrak{N})$ be the manifolds associated with $\overline{T}$ in the manner just indicated. Then $(\mathfrak{M}) = \mathfrak{N}^\perp$, $(\mathfrak{N}) = \mathfrak{M}^\perp$, $\mathfrak{M} = (\mathfrak{N})^\perp$, $\mathfrak{N} = (\mathfrak{M})^\perp$.*

By definition, $\overline{T}F = G$ with $F, G \in (\mathfrak{B})$, if $F(Tf) = Gf$ for all $f \in \mathfrak{B}$.‡ Let $F \in (\mathfrak{M})$; then $\overline{T}F = 0$ or $FTf = 0$, hence $F \perp \mathfrak{N}$, $F \in \mathfrak{N}^\perp$; thus $(\mathfrak{M}) \subset \mathfrak{N}^\perp$. Conversely, let $F \in \mathfrak{N}^\perp$; then $F(Tf) = 0$, $\overline{T}F = 0$, $F \in (\mathfrak{M})$ and $\mathfrak{N}^\perp \subset (\mathfrak{M})$. Thus $(\mathfrak{M}) = \mathfrak{N}^\perp$. Now let $F \in (\mathfrak{N})$; then for $\epsilon > 0$ there exists an F' with $\|F - \overline{T}F'\| \leqq \epsilon$. If $f \in \mathfrak{M}$, $\overline{T}F'f = F'Tf = 0$; thus $\overline{T}F' \perp \mathfrak{M}$, $\overline{T}F' \in \mathfrak{M}^\perp$. Since ϵ is arbitrary, $F \in \mathfrak{M}^\perp$ or $(\mathfrak{N}) \subset \mathfrak{M}^\perp$. Suppose $(\mathfrak{N}) \neq \mathfrak{M}^\perp$, then there exists an $F \in \mathfrak{M}^\perp$ such that $F \notin (\mathfrak{N})$. Since $\mathfrak{B}$ is reflexive, there exists an $f \in \mathfrak{B}$ such that $Ff = 1$, $f \perp (\mathfrak{N})$, or $\overline{T}Gf = GTf$

---

* B. J. Pettis, this Bulletin, abstract 44-7-295 (1938); Duke Mathematical Journal, vol. 5 (1939), pp. 249–253; D. Milman, *On some criteria for the regularity of spaces of the type (B)*, Comptes Rendus de l'Académie des Sciences de l' URSS, vol. 20 (1938), pp. 243–246. This author uses the terminology originally introduced by H. Hahn (*Über lineare Gleichungssysteme in linearen Raümen*, Journal für die reine und angewandte Mathematik, vol. 157 (1927), pp. 214–229) and calls a reflexive space "regular."

† Lorch, loc. cit.

‡ S. Banach, *Théorie des Opérations Linéaires*, p. 100.

$=0$ for all $G \in (\mathfrak{B})$. Hence $Tf=0, f \in \mathfrak{M}$, $Ff=0$ and not 1. This means that $(\mathfrak{N})=\mathfrak{M}^{\perp}$.

Since $\mathfrak{N}^{\perp}=(\mathfrak{M})$, $\mathfrak{N}^{\perp\perp}=\mathfrak{N}=(\mathfrak{M})^{\perp}$. Similarly, since $\mathfrak{M}^{\perp}=(\mathfrak{N})$, $\mathfrak{M}^{\perp\perp}=\mathfrak{M}=(\mathfrak{N})^{\perp}$.

THEOREM 2. *Let $\mathfrak{B}$ be a reflexive vector space. Let V be a bounded linear transformation in $\mathfrak{B}$ such that $|V^n| \leqq K$, $n=0, 1, 2, \cdots$. Then the sequence $\{T_n\}$ where $T_n=(1/n)\sum_{s=0}^{n-1} V^s$ converges strongly to a limiting transformation P which is a projection. The relation $Pf=f$ holds precisely for those elements f for which $Vf=f$. The relation $Pf=0$ holds precisely for the elements of the closed linear manifold spanned by elements of the form $Vg-g$, g arbitrary in $\mathfrak{B}$. The bound of P satisfies the inequality $|P| \leqq K$.*

Let $\mathfrak{M}$ be the c.l.m. of elements f for which $Vf=f$. Then $T_nf=f \to f$.

Let $\mathfrak{N}$ denote the c.l.m. spanned by elements of the form $Vg-g$, g arbitrary in $\mathfrak{B}$. If $f \in \mathfrak{N}$ and $\epsilon>0$ is arbitrary, then $g, h \in \mathfrak{B}$ may be found such that $f=Vg-g+h$ with $\|h\| \leqq \epsilon$. Now $T_nf=(1/n)(V^ng-g)+T_nh$. Furthermore $|T_n| \leqq (1/n)\sum_{s=0}^{n-1}|V^s| \leqq K$; therefore $\|T_nf\| \leqq (2/n)K\|g\|+K\epsilon \leqq 2K\epsilon$ providing n is sufficiently large. Thus $T_nf \to 0$ when $f \in \mathfrak{N}$.

Now let $f \in \mathfrak{M}$, $g \in \mathfrak{N}$; then $\|f\| \leqq K\|f+g\|$. This follows from $|T_n| \leqq K$ and $T_n(f+g) \to f$. As a consequence the only element common to $\mathfrak{M}$ and $\mathfrak{N}$ is 0. In addition, the totality of elements of the form $f+g$ is a c.l.m. which will be designated by $\mathfrak{M} \dot{+} \mathfrak{N}$. The strong limit of T_n exists in this c.l.m. and is a b.l.t. P of bound $|P| \leqq K$. Clearly $P^2=P$ in $\mathfrak{M} \dot{+} \mathfrak{N}$. To complete the proof, it must be shown that $\mathfrak{M} \dot{+} \mathfrak{N}=\mathfrak{B}$. This equation will be established with the aid of the assumption, not employed up to this point, that $\mathfrak{B}$ is reflexive.

If V is the adjoint of $\overline{V}$, then $\overline{V}^n=\overline{(V^n)}$ and $|\overline{V}^n| \leqq K$. Thus what has just been established for V may be elaborated for $\overline{V}$. Let $(\mathfrak{M})$ and $(\mathfrak{N})$ denote the manifolds found for $\overline{V}$ by the method just applied to V. Applying Theorem 1 to the transformations $V-I$ and $\overline{V}-\overline{I}=\overline{(V-I)}$, we have $(\mathfrak{M})=\mathfrak{N}^{\perp}$, $(\mathfrak{N})=\mathfrak{M}^{\perp}$. If $f \in \mathfrak{B}$ but $f \notin \mathfrak{M} \dot{+} \mathfrak{N}$, there exists an $F \in (\mathfrak{B})$ such that $Ff=1$, $Fg=0$ for $g \in \mathfrak{M} \dot{+} \mathfrak{N}$. Hence $F \perp \mathfrak{M}$, $F \perp \mathfrak{N}$, or $F \in \mathfrak{M}^{\perp}=(\mathfrak{N})$, $F \in \mathfrak{N}^{\perp}=(\mathfrak{M})$. As $(\mathfrak{M})$ and $(\mathfrak{N})$ have only the element 0 in common, $F=0$, which contradicts $Ff=1$. Thus $\mathfrak{M} \dot{+} \mathfrak{N}=\mathfrak{B}$.

BARNARD COLLEGE, COLUMBIA UNIVERSITY

ON A CERTAIN CLASS OF SYMMETRIC HYPERSURFACES

DARRELL R. SHREVE

In the literature of algebraic geometry, relatively little is written on hypersurfaces H of order n and of $r-1$ dimensions in S_r, invariant under the symmetric permutation group G on $r+2$ homogeneous coordinates whose sum is zero. Surfaces in S_3 invariant under the symmetric G_{120} have been studied by Emch [1]; the Clebsch diagonal surface has been discussed by Clebsch [2], Eckardt [3], and Ciani [4, 5, 6]. The Segre cubic variety in S_4 has been investigated by Segre [8]; Ciani [7] has developed properties of loci in S_4 invariant under the G_{720}.

It is well known that the equation of such a hypersurface H is expressible uniquely in terms of the elementary symmetric functions p_i (of order i in $x_1, \cdots, x_{r+2}$), or in terms of the Σ functions $\Sigma_n=\sum_{i=1}^{r+2}x_i^n$. With the condition $\Sigma_1=0$, any H is a member of the linear system

$$\sum_{i=1}^{N} C_i\Sigma_2^a\Sigma_3^b\cdots\Sigma_{r+2}^d = 0 \tag{1}$$

where N is the number of nonnegative solutions of the Diophantine equation

$$2a + 3b + \cdots + (r+2)d = 0. \tag{2}$$

It follows immediately that there is a unique hyperquadric H_2, $\Sigma_2=0$ or $p_2=0$, a unique cubic H_3, $\Sigma_3=0$ or $p_3=0$, a pencil of quartics H_4, $\lambda\Sigma_4+\mu\Sigma_2^2=0$, a pencil of quintics H_5, $\lambda\Sigma_5+\mu\Sigma_2\Sigma_3=0$, and so on, in S_r invariant under G.

Since there are no real points on $\Sigma_n=0$ if n is even, values of n will be restricted to odd positive integers throughout the remainder of this paper.

Emch [1] has shown that the equation of any surface of odd order in S_3, invariant under the symmetric G_{120} on 5 x's whose sum is zero necessarily has the form $Ap_3+Bp_5=0$, which is equivalent to $\lambda\Sigma_3+\mu\Sigma_5=0$. The obvious generalization of this statement is that any H of order n necessarily is a member of the linear system

$$A\Sigma_3 + B\Sigma_5 + \cdots + D\Sigma_{r+2} = 0; \tag{3}$$

if $r=2k+1$, and in case $r=2k$, of the system

$$A\Sigma_3 + B\Sigma_5 + \cdots + C\Sigma_{r+1} = 0. \tag{4}$$

If $r=2k+1$, all H of (3) contain the base $\Sigma_3=\Sigma_5=\cdots=\Sigma_{r+2}=0$. This base is of order $\alpha=(r+2)\cdot r!/2^k\cdot k!$ and of dimension k, consisting of the α subspaces into which the S_k $[x_1+x_2=0,\ x_3+x_4=0,\ \cdots,\ x_r+x_{r+1}=0,\ x_{r+2}=0]$ is carried by the substitutions of G. Each S_k is invariant under a subgroup of G of order $2^{k+1}\cdot(k+1)!$.

If $r=2k$, all H of (4) contain the base $\Sigma_3=\Sigma_5=\cdots=\Sigma_{r+1}=0$, of order $\beta=(r+1)!/2^k\cdot k!$ and of dimension k, consisting of the β subspaces into which the substitutions of G carry the S_k $[x_1+x_2=0,\ x_3+x_4=0,\ \cdots,\ x_{r+1}+x_{r+2}=0]$; each S_k is invariant under a subgroup of G of order $2^{k+1}\cdot(k+1)!$.

The hypersurfaces $\Sigma_n=0$ have as double points only those points whose coordinates are proportional to $(n-1)$th roots of unity, since at a double point the $r+1$ partial derivatives $\partial\Sigma_n/\partial x_i=nx_i^{n-1}-nx_{r+2}^{n-1}$, $i=1,\ 2,\ \cdots,\ r+1$, must vanish (with the dependence of x_{r+2} expressed by $\Sigma_1=0$). It follows immediately that $\Sigma_n=0$ has $C_{r+1,k}$ real double points if $r=2k$, and has no real double points if $r=2k+1$. Imaginary double points of $\Sigma_n=0$ will occur whenever there is a set of $r+2$ $(n-1)$th roots of unity, not all real, whose sum is zero.

If $r=2k+1$, then none of the α subspaces S_k passes through a double point of $\Sigma_n=0$, since at a double point no $x_i=0$.

If $r=2k$, then in each of the β subspaces S_k there are $(n-1)^k$ double points of $\Sigma_n=0$, of which exactly 2^k are real; through each of the real double points pass $(k+1)!$ subspaces S_k. If $r=2k$, and $2k+1$ is a prime, then the β subspaces S_k divide into $(r-1)!$ sets of $(r+1)$-tuples, each $(r+1)$-tuple being transformable into another $(r+1)$-tuple by a cyclic substitution of G of period $r+1$. (This is a generalization of the 6 quintuples of planes on the Segre cubic variety.)

The hypersurfaces $\Sigma_n=0$ can contain no points of multiplicity greater than two, since not all the partial derivatives of higher order vanish at any point.

Eckardt [3] has given an admirable synthetic and analytic discussion of the properties of an Eckardt point of a surface in S_3. The analytic generalization is immediate. Let a generalized Eckardt point E of a hypersurface F, of order $m>2$ and of $r-1$ dimensions in S_r, be a simple point of F such that the hyperplane T tangent to F in E intersects F in a hypercone with a vertex at E. We may so choose the polylateral of reference in S_r that E is $(1,\ 0,\ \cdots,\ 0)$, and T is $x_2=0$. Then the equation of F necessarily has the form

$$x_2x_1^{m-1}+a_1x_2x_1^{m-2}+\cdots+a_{m-2}x_2x_1+a_m=0 \tag{5}$$

where a_i is a form of order i in $x_2,\ x_3,\ \cdots,\ x_{r+1}$.

It follows immediately that the ith polar of E reduces into T

$(x_2=0)$ and a hypersurface of order $m-i-1$ which does not pass through E.

Conversely, if the polar of a point P with respect to a hypersurface F of order m reduces into a hyperplane π passing through P and a hypersurface of order $m-2$ not passing through P, then if F does not reduce into π and a hypersurface of order $m-1$, P is a generalized Eckardt point of F and π is tangent to F at the point P.

On each hypersurface $\Sigma_n=0$ in S_r there are $C_{r+2,2}$ Eckardt points E_{ij} $(x_i=-x_j\neq 0;\ x_s=0,\ s\neq i, j)$, with the hyperplane $x_i+x_j=0$ tangent to $\Sigma_n=0$ at E_{ij}. The polar of E_{ij} with respect to $\Sigma_n=0$ is $x_i^{n-1}-x_j^{n-1}=0$, which contains the hyperplanes $x_i+x_j=0$ and $x_i-x_j=0$. The latter is the axis of the perspectivity (ij), with center at E_{ij}, under which $\Sigma_n=0$ and the polar of E_{ij} are invariant. The $(n-1)$-fold locus $x_i=x_j=0$ of the polar of E_{ij} contains $C_{r,2}$ Eckardt points E_{st}, $(s, t\neq i, j)$.

If $r=2k$, the hyperplane $x_i+x_j=0$ contains $C_{r,k}$ real double points of $\Sigma_n=0$, and $x_i-x_j=0$ contains the remaining $C_{r,k-1}$ real double points. Each real double point D of $\Sigma_n=0$ is collinear with $(k+1)^2$ couples of points of $\Sigma_n=0$, each couple consisting of an E_{ij} and the double point corresponding to D under (ij). In each of the β subspaces S_k on $\Sigma_n=0$ there are $k+1$ points E_{ij}, and through each E_{ij} there pass $r!/2^k\cdot k!$ subspaces S_k, while the three collinear points E_{ij}, E_{ik}, E_{jk} do not lie in a common S_k on $\Sigma_n=0$. An S_k is the locus of points common to the $k+1$ hyperplanes tangent to Σ_n at the $k+1$ points E_{ij} in the S_k.

If $r=2k+1$, through each E_{ij} on $\Sigma_n=0$ pass $(r+1)!/2^k\cdot k!$ subspaces S_k, with $k+1$ points E_{ij} in each S_k.

The Eckardt point of a hypersurface F is in general of multiplicity $r-1$ on the Hessian of F; the Eckardt points of $\Sigma_n=0$ are of multiplicity $(r-1)(n-2)$ on the Hessian of $\Sigma_n=0$, whose equation is $\sum_{i=1}^{r+2}1/x_i^{n-2}=0$.

References

1. Arnold Emch, *Some geometric applications of the symmetric collineation groups on five variables*, Tôhoku Mathematical Journal, vol. 40 (1934), pp. 216–221.

2. A. Clebsch, *Ueber die Anwendung der quadratischen Substitution auf die Gleichungen 5ten Grades und die geometrische Theorie des ebenen Fünfseits*, Mathematische Annalen, vol. 4 (1871). See in particular pp. 331–340.

3. F. E. Eckardt, *Ueber diejenigen Flächen dritten Grades, auf denen sich drei gerade Linien in einem Punkte schneiden*, Mathematische Annalen, vol. 10 (1876), pp. 227–272.

4. Edgardo Ciani, *Sulla superficie diagonale di Clebsch*, Rendiconti della Accademia dei Lincei, vol. 7, no. 6.

5. ———, *Sul pentaedro complete*, Rendiconti della Accademia dei Lincei, vol. 7, no. 5.

6. ———, *Sopra la configurazione del pentaedro*, Rendiconti del Circolo Matematico di Palermo, vol. 21 (1906), pp. 322–341.

7. ———, *Una interpretazione geometrica del gruppo totale di sostituzioni sopra sei elementi*, Annali di Matematica Pura ed Applicata, 1909.

8. Corrado Segre, *Sulla varietà cubica con dieci punti doppi dello spazio a quattro dimensioni*, Atti della Accademia delle Scienze di Torino, vol. 22 (1886–1887), p. 791.

PURDUE UNIVERSITY

CORRECTION TO "ON GREEN'S FUNCTIONS IN THE THEORY OF HEAT CONDUCTION IN SPHERICAL COORDINATES"*

A. N. LOWAN

In the article entitled *On Green's functions in the theory of heat conduction* by H. S. Carslaw and J. C. Jaeger (this Bulletin, vol. 45 (1939), pp. 407–413), a misprint is noted in the expression for G on page 133 of my article *On the operational determination of two dimensional Green's functions in the theory of heat conduction* (this Bulletin, vol. 44 (1938), pp. 125–133), the correct expression for G being

$$G = u + v = \frac{1}{4\pi} \sum_{n=-\infty}^{\infty} \cos n(\theta - \theta_0) \int_{-\infty}^{\infty} \alpha e^{-k\alpha^2 t} \frac{H_n^{(1)}(\alpha r_0)}{U_n(a\alpha)}$$
$$\cdot \left\{ J_n(\alpha r) U_n(\alpha a) - H_n^{(1)}(\alpha r) \left[\alpha \frac{d}{dz} J_n(z) + h J_n(z) \right]_{z=\alpha a} \right\} d\alpha,$$

where

$$U_n(\alpha a) = \left[\alpha \frac{d}{dz} H_n^{(1)}(z) + h H_n^{(1)}(z) \right]_{z=\alpha a}.$$

When this correct expression is employed, formula (20), page 313, of the present paper becomes

$$G(r, \theta, \phi, t; r_0, \theta_0, \phi_0) = \frac{1}{8\pi (r r_0)^{1/2}} \sum_{n=0}^{\infty} (2n+1) P_n(\cos \gamma)$$

(A)
$$\cdot \int_{-\infty}^{\infty} \alpha e^{-k\alpha^2 t} \frac{H_{n+1/2}^{(1)}(\alpha r_0)}{U_{n+1/2}(\alpha a)} \left\{ J_{n+1/2}(\alpha r) U_{n+1/2}(\alpha a) - H_{n+1/2}^{(1)}(\alpha r) \right.$$
$$\left. \cdot \left[\alpha \frac{d}{dz} J_{n+1/2}(z) + (h - 1/(2a)) J_{n+1/2}(z) \right]_{z=a\alpha} \right\} d\alpha.$$

* This Bulletin, vol. 45 (1939), pp. 310–315.

Further, when the last expression is integrated in accordance with the statement in the last paragraph of page 314, the correct version of formula (26) becomes

$$G(r, t; r_0) = \frac{1}{8\pi^{3/2}(kt)^{1/2}rr'}\left[\exp\left\{-\frac{(r-r')^2}{4kt}\right\}\right.$$
$$+ \exp\left\{-\frac{(r+r'-2a)^2}{4kt}\right\} + 2\left(h - \frac{1}{a}\right)$$
$$\left.\cdot \int_0^\infty \exp\left\{\left(h - \frac{1}{a}\right)\xi - \frac{(r+r'+\xi-2a)^2}{4kt}\right\} d\xi\right].$$

The above developments in the second paragraph of page 315 of the present paper ending with equation (28) are incorrect, the correct version of formula (28) being once more (A).

The factor $J_n(\alpha r)$ in the bracket of equation (22) on page 314 should be replaced by $J_{n+1/2}(\alpha r)$.

The factor 1/2 in equation (1) on page 310, should be replaced by 1/8.

The phrase "the former analogy" in the first paragraph of the second section of page 313 should read "the formal analogy."

I am indebted to Professor H. S. Carslaw, for calling my attention to the errors above mentioned and for suggesting the appropriate corrections.

BROOKLYN COLLEGE AND
YESHIVA COLLEGE